

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:46:52 ; Search time 29 seconds
(without alignments)
142.101 Million cell updates/sec

Title: US-09-926-234-1
Perfect score: 117
Sequence: 1 HGVHGFGRGFGFGFGPSC 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	68.4	270	5 Q9VYMS	Q9VYMS drosophila
2	73	62.4	107	10 Q9LEP7	Q9LEP7 brassica na
3	73	62.4	375	6 Q9NRA1	Q9NRA1 macaca fasc
4	72	63.4	486	10 Q9AP23	Q9AP23 oryza sativ
5	72	61.5	426	4 Q12937	Q12937 homo sapien
6	72	61.5	501	4 Q9WMB4	Q9WMB4 homo sapien
7	72	61.5	521	4 Q9Q932	Q9Q932 homo sapien
8	72	61.5	533	16 Q69975	Q69975 streptomyces
9	71	60.7	253	5 Q9WMB7	Q9WMB7 heliocidari
10	71	60.7	254	5 Q9WMB6	Q9WMB6 heliocidari
11	71	60.7	1260	5 Q9WMB5	Q9WMB5 drosophila
12	71	60.7	1379	5 Q9WMB4	Q9WMB4 drosophila
13	70	59.8	229	5 Q9XUI2	Q9XUI2 caenorhabdi
14	70	59.8	259	12 Q98418	Q98418 parametium
15	70	59.8	519	16 Q93J30	Q93J30 streptomyces
16	69	59.0	268	16 Q93Z60	Q93Z60 arabidopsis

17	69	59.0	416	10 Q9LDB4	Q9LDB4 arabidopsis
18	69	59.0	432	10 Q99075	Q99075 arabidopsis
19	69	59.0	1257	4 Q14554	Q14554 homo sapien
20	68	58.1	137	10 Q94731	Q94731 arabidopsis
21	68	58.1	148	10 Q940R3	Q940R3 arabidopsis
22	68	58.1	185	5 Q18507	Q18507 tenebrio mo
23	68	58.1	185	5 Q27016	Q27016 tenebrio mo
24	68	58.1	617	3 Q9P6C4	Q9P6C4 neurospora
25	67	57.3	629	5 Q8T1Y2	Q8T1Y2 dictyosteli
26	66	56.4	186	13 Q9DEX8	Q9DEX8 cyprinus ca
27	66	56.4	215	6 Q28130	Q28130 bos taurus
28	66	56.4	658	11 Q9UL61	Q9UL61 mus musculu
29	65	55.6	396	10 Q65450	Q65450 arabidopsis
30	65	55.6	790	5 Q22872	Q22872 caenorhabdi
31	65	55.6	1203	5 Q21815	Q21815 caenorhabdi
32	64	54.7	100	12 Q89403	Q89403 parametium
33	64	54.7	147	5 Q9VY30	Q9VY30 drosophila
34	64	54.7	372	5 Q9VNB8	Q9VNB8 drosophila
35	64	54.7	421	13 Q9DEY1	Q9DEY1 cyprinus ca
36	63	53.8	47	10 Q9S857	Q9S857 glycine max
37	63	53.8	140	5 Q9NL39	Q9NL39 pinus taeda
38	63	53.8	205	10 Q22432	Q22432 pinus taeda
39	63	53.8	687	5 Q9W3D6	Q9W3D6 drosophila
40	63	53.8	2703	5 Q9VEG7	Q9VEG7 drosophila
41	63	53.8	2715	5 Q61603	Q61603 drosophila
42	62	53.0	91	2 P77018	P77018 escherichia
43	62	53.0	210	10 Q41187	Q41187 arabidopsis
44	62	53.0	215	5 Q9G0H3	Q9G0H3 drosophila
45	62	53.0	217	5 Q9G0H1	Q9G0H1 drosophila

ALIGNMENTS

ID	Q9VYMS	PRELIMINARY:	PRT:	270 AA.
AC	Q9VYMS	01-MAY-2000 (TRENBLrel. 13, Created)		
DT	01-MAY-2000 (TRENBLrel. 13, Last sequence update)			
DT	01-MAY-2000 (TRENBLrel. 21, Last annotation update)			
DE	CG2560	protein (R57452P).		
GN	CG2560			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George P.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Burton G.G., Wortman J.P., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Sutton P.C., Rogers Y.-H.C., Blazer P.G., Champe M., Pfeiffer B.D.,			
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.P., Myles G.L.G.,			
RA	Abt J.J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,			
RA	Ballew P.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.K.,			
RA	Beeson K.Y., Benos P.V., Bertram B.P., Bhandari P., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cardieu E., Conter A., Chandra J.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.P., Davies P.,			
RA	Deeb S., Delcher A., Deng Z., Dey A.D., Dew J., Dietz S.M.,			
RA	Dodson K., Dunn L.E., Dunham M., Dugan-Rocha S., Dunkov R.C., Dunn P.,			
RA	Durkin K.D., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hosr D., Houston K.A., Howland T.J., Wei M.-H., Ijames C.,			
RA	Jalili M., Kalush F., Karpen G.H., Ke Z., Kennisn J.A., Ketchum K.A.,			
RA	Kimmel R.E., Kodira C.P., Kraft C., Pravitz S., Pulp D., Lai Z.,			

BEST AVAILABLE COPY

RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Moruklov G., Mlshina N.V., Mobary C., Morris J., Moshieff A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.P., Nelson K.A., Nixon K., Nuskern D.P., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard T., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders P.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas P., Tector C., Turner P., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Wesselsbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhu W., Zhu X., Zhu G., Zhu Q., Smith H.G.,
 RA Gibbs P.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).

PP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champagne M., Chavez G., Dorsett V., Drensek D., Farfan D., Frise E.,
 RA George P., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Pargass V., Park S.,
 RA Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celinker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO03489; AAF48167.1; -;
 DR EMBL: AY011531; AAJ49153.1; -;
 DR FLYbase: FBgn0030394; GQ2560.
 DR InterPro: IPR000618; Insect_cuticle.
 DR Pfam: PF00379; Cuticlin_bind_4; 1.
 DR PROSITE: PS00233; CUTICLE; 1.
 SQ SEQUENCE 270 AA; 24715 MW; C42982E27P1AB6B CPC64;

Query Match 68.4%; Score 80; DB 5; Length 270;
 Best Local Similarity 70.0%; Pred. No. 0.014;
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVGPCTGPGSG 20
 Db 210 NGVGHGFGNGVGPCTGPGSG 229

RESULT 2
 Q9LEP7 PRELIMINARY; PRT; 107 AA
 ID Q9LEP7;
 AC Q9LEP7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Putative glycine-rich protein.
 DE Brassica napus (Rape).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Poastiales;
 CC Eucosids; II; Brassicales; Brassicaceae; Brassica.
 CX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV N-0-9;
 RA Bowers N.L., Trick M.;
 RT "Microsynteny at the FCA region between *Arabidopsis thaliana* and
 RT *Brassica napus*.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ293726; CAC01931.1; -;
 DR InterPro: IPR001525; CS_TNA_meth.
 DR PROSITE: PS00095; CS_MTASE_2; UNKNOWN 1.
 SQ SEQUENCE 107 AA; 9714 MW; 0A7BFAA01P0C1B3 CPC64;

Query Match 62.4%; Score 73; DB 10; Length 107;
 Best Local Similarity 60.0%; Pred. No. 0.04;
 Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVGPCTGPGSG 20
 Db 78 NGVGHGFGNGVGPCTGPGSG 97

RESULT 3
 Q9NOA1 PRELIMINARY; PRT; 375 AA
 ID Q9NOA1;
 AC Q9NOA1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Unnamed protein product.
 DE Macaca fascicularis (Crab eating monkey) (Cynomys) (monkey).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecoidea; Macaca.
 CX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oada N., Hida M., Fuseda T., Tanuma P., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB046029; BAB01611.1; -;
 SQ SEQUENCE 375 AA; 39100 MW; 77E9A447E2P166A7 CPC64;

Query Match 62.4%; Score 73; DB 6; Length 375;
 Best Local Similarity 65.0%; Pred. No. 0.15;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVGPCTGPGSG 20
 Db 307 SGVAGAGVAGAGPAGPAGAG 326

RESULT 4
 Q9AR23 PRELIMINARY; PRT; 486 AA
 ID Q9AR23;
 AC Q9AR23;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE Putative glycine-rich protein.
 DE Oryza sativa (rice).
 GN O01174.D05.13 OR P066604.1.
 OS Oryza sativa (rice)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzoideae; Oryza.
 CX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV NIPONBARE;
 PA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa niponbare (GAJ) genomic DNA, chromosome 1, BAC
 RT clone O01174.D05.13.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV NIPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa niponbare (GAJ) genomic DNA, chromosome 1, PAC
 RT clone P066604.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003118; BAB33021.1; -;
 DR EMBL: AP003047; BAB32899.1; -;
 DR InterPro: IPR002952; Eggshell.
 DR PRINTS: PP01228; EGGSHL.
 SQ SEQUENCE 486 AA; 38006 MW; 4B533FEF03FD667 CPC64;

Query Match 62.4%; Score 73; DB 10; Length 486;
 Best Local Similarity 73.7%; Pred. No. 0.19;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0.

OY 2 GVGHGFGNGVGPCTG--PSSG 20
 DB 339 GVGNGFGVGAGAGTGLSG 357

RESULT 5

O12937

PRELIMINARY;

PRT; 426 AA.

AC Q12937;

01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE BS-84.

HSD-1.

GN HSD-1.

HSD-1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

NCBI_TaxID=9606;

OX NCBI_TaxID=9606;

RN [1]

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RA Miao S.;

RT "Studies on cDNA encoding a human sperm membrane protein BS-84";

PL Prog Nat Sci 5:119-122(1995).

RN [2]

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RA Miao S.;

RT "Studies on cDNA encoding a human sperm membrane protein BS-84";

PL Prog Nat Sci 5:119-122(1995).

RN [2]

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

DR EMBL: BC019247; AAH19247.1; -.
 DR InterPro: IPR000104; Antifreeze_1.
 DP PRINTS: PR00308; ANTIFREEZE1
 SQ SEQUENCE 501 AA; 53451 MW; 63DBD632579A6D84 CRC64;

Query Match 61.5%; Score 72; DB 4; Length 501;
 Best Local Similarity 63.6%; Pred. No. 0.26;
 Matches 14; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

OY 1 NGVGHGFGNGVGPCTG--PSSG 20

DB 151 SGAGHGSGSGSGGCGCGVPSG 172

RESULT 7

O99932

PRELIMINARY;

PRT; 523 AA.

AC O99932;

01-MAY-1997 (TReMBLrel. 03, Created)

DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)

DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE HSM-1.

HSM-1.

GN HSM-1.

HSM-1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

NCBI_TaxID=9606;

OX NCBI_TaxID=9606;

RN [1]

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RX MEDLINE=96380169; PubMed=8788182;

RA Liu Q.Y., Wang L.F., Miao S.Y., Catterall J.F.;

RT "Expression and characterization of a novel human sperm membrane protein";

PL Biol Reprod 54:323-330(1996).

RN [3]

RN [3]

RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinaishi H., Hopwood D.A.,
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome,"
 RL Mol. Microbiol. 21:77-96(1996).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bertley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.P., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandram M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
 RA Warren T., Wierzyk A., Woodward J., Barrall B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL: AL022268; CA018336.1; -;
 DR HSSP: P16932; 2DKR;
 DR InterPro: IPR000954; AminoTran_3;
 DR InterPro: IPR001092; HLH_basic;
 DR Pfam: PF00202; aminotran_3; 1;
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1;
 KW AminoTransferase; Transferase;
 SQ SEQUENCE 532 AA; 53554 MW; B6749ECL5B162315 CRC64;

Query Match 61.5%; Score 72; DB 16; Length 532;
 Best Local Similarity 68.4%; Pred. No. 0.28;
 Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 2 GVGHGNGVGPCTGPGSG 20
 DB 84 GGTGPGTGTGPGTGTG 102

RESULT 9
 ID 095W97 PRELIMINARY; PRT; 253 AA.
 AC 095W97;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Primary mesenchyme-specific protein PM27.
 GN PM27.
 OS Helicoidaris tuberculata (Sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Echinometridae;
 OC Helicoidaris;
 OX NCBI_TaxId=7635;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Stander M.C., Raff P.A.;
 RT "Gene expression and skeleton formation in the sea urchin Helicoidaris
 RT erythrogramma,"
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF359342; AAL27535.1;
 DR InterPro: IPR001304; Lectin_C;
 DR Pfam: PF00059; lectin_c; 1;
 DR PROSITE: PS50041; C TYPE LECTIN 2; 1;
 SQ SEQUENCE 253 AA; 27331 MW; 555417A0CE0778F4 CRC64;

Query Match 60.7%; Score 71; DB 5; Length 253;
 Best Local Similarity 63.2%; Pred. No. 0.17;
 Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 2 GVGHGNGVGPCTGPGSG 20
 DB 30 GMSPGMGPGMGPGMGPG 48

RESULT 10
 ID 095W96 PRELIMINARY; PRT; 254 AA.
 AC 095W96;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Primary mesenchyme-specific protein PM27.
 GN PM27.
 OS Helicoidaris erythrogramma (Sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Echinometridae;
 OC Helicoidaris;
 OX NCBI_TaxId=7634;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Stander M.C., Raff P.A.;
 RT "Gene expression and skeleton formation in the sea urchin Helicoidaris
 RT erythrogramma,"
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF359343; AAL27536.1;
 DR InterPro: IPR001304; Lectin_C;
 DR Pfam: PF00059; lectin_c; 1;
 DR PROSITE: PS50041; C TYPE LECTIN 2; 1;
 SQ SEQUENCE 254 AA; 27430 MW; 48246B97966EE554 CRC64;

Query Match 60.7%; Score 71; DB 5; Length 254;
 Best Local Similarity 63.2%; Pred. No. 0.17;
 Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 2 GVGHGNGVGPCTGPGSG 20
 DB 30 GMSPGMGPGMGPGMGPG 48

RESULT 11
 ID 09W2Y5 PRELIMINARY; PRT; 1260 AA.
 AC 09W2Y5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE CG2971 protein.
 GN CG2971 OR CG12646.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hopkins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.P., Yandell M.D., Zhang Q., Chen L.Y.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfiffner B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrait J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borkov D., Borzian M.R., Bouck J., Brockstein P., Brotler P.,
 RA Burtis K.C., Burnas D.A., Butler H., Cantu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.F., Houck J.,
 RA Hostin D., Houston K.A., Howland J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Mostrelet A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskens D.P., Paclet J.M.,
 RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert F., Remington K., Sanders P.D., Schaefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkac P., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Wordage T., Worley K.C., Wu P., Yang S., Yao Q.A.,
 RA Ye J., Yen R.F., Zavertil J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.C.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003449; AAF4552.1;
 DR Flybase: FBgn0030176; CG2971.
 SO SEQUENCE 1260 AA; 134534 MW; AAFDDIFQDF0F3A1 CRC64;

Query Match 60.7%; Score 71; DB 5; Length 1260;
 Best Local Similarity 70.0%; Pred. No. 0.9;
 Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVPGTGPSSG 20

DB 1019 NGPPGJFGNGPGGPGSSG 1038

RESULT 12

Q8SWX1 PRELIMINARY; PRT; 1379 AA.

AC Q8SWX1;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE LD11664p (Fragment).
 GN CG2971.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_Taxid=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuno U., Paclet J., Paragae V., Park S.,
 RA Patel S., Phuenhavanong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY095027; AAM1355.1;
 DR NON-TER 1379 1379
 SO SEQUENCE 1379 AA; 147165 MW; AD1533BDF4F49986 CRC64;

Query Match 60.7%; Score 71; DB 5; Length 1379;
 Best Local Similarity 70.0%; Pred. No. 0.9;
 Matches 14; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVPGTGPSSG 20

DB 1019 NGPPGJFGNGPGGPGSSG 1038

RESULT 13

Q9XU12

AC Q9XU12; PRELIMINARY; PRT; 229 AA.

DT 01-NOV-1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ZK39.2 protein.
 GN ZK39.2.
 OS *Caenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Phadidida; Phadidiroidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_Taxid=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Keshaw J.K.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN (2)
 PP SEQUENCE FROM N.A.
 PY MEDLINE=9906613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology,"
 RL Science 282:2012-2018(1998).
 DR EMBL: Z82093; CAB05018.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2, 1.
 SO SEQUENCE 229 AA; 23895 MW; D90DD4SF5A4EC724 CRC64;

Query Match 59.8%; Score 70; DB 5; Length 229;
 Best Local Similarity 65.0%; Pred. No. 0.21;
 Matches 13; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVPGTGPSSG 20

DB 30 NGHGHGNGNGNGNGNG 49

RESULT 14

Q98438 PRELIMINARY; PRT; 259 AA.

AC Q98438;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE A386R protein.
 GN A386R.
 OS *Parametium bursaria* chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 OC NCBI_Taxid=10506;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96400190; PubMed=8806566;
 RA Kutsch G.F., Li Y., Lu Z., Furuta M., Rock D.L., Van Etten J.L.;
 RT "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map
 RT positions 182 to 258,"
 RL Virology 223:303-317(1996).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20013326; PubMed=10544099;
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
 RA Lisee A.D., Nickerson K.W., Van Etten J.L.;
 RT "Chlorella virus PBCV-1 encodes a functional homospertmidine
 RT synthase,"
 RL Virology 263:254-262(1999).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20478054; PubMed=11021991;
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
 RT PBCV-1,"
 RL Virology 276:27-36(2000).
 RN (4)
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

RN [5]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Graves M.V., Van Etten J.L.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RA Gurnon J.R., Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U42580; AAC96754.1; -;
 SO SEQUENCE 259 AA; 25950 MW; F303D2E495815F CRC64;

Query Match 59.8%; Score 70; DB 12; Length 259;
 Best Local Similarity 70.6%; Pred. No. 0.24;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 4 GHGFGNGVGPPTGPGSG 20
 DB 86 GAGFGVGLGPGPGAG 102

RESULT 15
 093330
 ID 093330 PRELIMINARY; PRT; 519 AA.
 AC 093330;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Putative protease.
 GN SC03977 OR SCBAC25E3.14.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Collins M., Harris D.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Ribbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Metzorrore A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomyete Streptomyces
 coelicolor A3(2)";
 RL Nature 417:141-147(2002).
 DR EMBL; AL596251; CAC44701.1; -;
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR Hydrolase; Protease; Serine protease.
 SO SEQUENCE 519 AA; 50327 MW; 5CB8D5F0CC19E428 CRC64;

Query Match 59.8%; Score 70; DB 16; Length 519;
 Best Local Similarity 60.0%; Pred. No. 0.48;
 Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 NGVGHFGNGVGPPTGPGSG 20
 DB 110 NGTNGFGDGPBGAGGAG 129

Search completed: May 6, 2003, 14:49:50
 Job time : 32 sec

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:45:57 ; Search time 11 seconds

75,412 Million cell updates/sec

Title: US-09-926-234 1

Perfect score: 117

Sequence: 1 NCVGHGPGNGVPGTGPSSG 20

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	60.7	267	PM27_STRPU	Q26616 strongyloce
2	68	58.1	347	MSA2_PLAF2	Q03646 plasmodium
3	67	57.1	344	GRPI_PETHY	P09789 peritremia h
4	66	56.4	5263	PBCH_BOMMO	P05790 bombyx mori
5	64	54.7	672	PHX5_MOUSE	P08199 mus musculu
6	63	53.8	104	HOL3_HOUDI	Q25055 holotrichia
7	61	52.1	158	PGSG_HUMAN	P10124 homo sapien
8	61	52.1	338	GRP_APATH	P27483 arabidopsis
9	60	51.3	88	PER_DPOTE	Q26287 drosophila
10	60	51.3	110	PER_DPOTE	Q26288 drosophila
11	60	51.3	114	PER_DPOTE	Q26289 drosophila
12	60	51.3	125	PER_DPOTE	Q03293 drosophila
13	60	51.3	141	PER_DPOTE	Q04537 drosophila
14	60	51.3	174	PER_DPOTE	P12347 xenopus lae
15	60	51.3	174	PER_DPOTE	P05781 acetabulari
16	60	51.3	174	PER_DPOTE	Q03355 drosophila
17	60	51.3	174	PER_DPOTE	Q03356 drosophila
18	60	51.3	174	PER_DPOTE	Q03357 drosophila
19	60	51.3	174	PER_DPOTE	Q03358 drosophila
20	60	51.3	174	PER_DPOTE	Q03359 drosophila
21	59	50.4	143	QID3_TRIHA	P07663 drosophila
22	59	50.4	143	QID3_TRIHA	P52129 bacterioph
23	59	50.4	143	QID3_TRIHA	P25129 bacterioph
24	59	50.4	143	QID3_TRIHA	P25129 bacterioph
25	58	49.6	152	PGSG_MOUSE	P13609 mus musculu
26	58	49.6	179	PGSG_MOUSE	P04317 rattus norv
27	58	49.6	179	PGSG_MOUSE	P13617 rattus norv
28	58	49.6	179	PGSG_MOUSE	P13617 rattus norv
29	58	49.6	179	PGSG_MOUSE	P13617 rattus norv
30	58	49.6	179	PGSG_MOUSE	P13617 rattus norv
31	57.5	49.1	1596	MAM_DROME	P21319 drosophila
32	57.5	49.1	1596	MAM_DROME	P21319 drosophila
33	57	48.7	458	KICM_HUMAN	P13646 homo sapien

ALIGNMENTS

RESULT 1	PM27_STRPU	STANDARD:	PRT:	267 AA.
AC	Q26616			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	27 kDa primary mesenchyme-specific spicule protein precursor.			
GN	PM27.			
OS	Strongylocentrotus purpuratus (Purple sea urchin).			
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;			
OC	Echinoidae; Euechinozoa; Echinodermata; Eleutherozoa; Echinozoa;			
OC	Strongylocentrotidae.			
OX	NCBI_Taxid=7668;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9526928, PubMed=7537234;			
RA	Harkey M.A., Klueng K., Sheppard P., Raff P.A.;			
RT	"Structure, expression, and extracellular targeting of PM27, a			
RT	skeletal protein associated specifically with growth of the sea			
RT	urchin larval spicule."			
RL	Dev. Biol. 168:549-566(1995).			
CC	- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF EXECUTION OF			
CC	SKELETAL GROWTH.			
CC	- SUBCELLULAR LOCATION: SECRETED TO THE SKELETAL COMPARTMENT AND			
CC	ACCUMULATES PREDOMINANTLY AT THE ADVANCING MINERALIZING SURFACE OF			
CC	THE SPICULE TIPS. AS THE SPICULES ELONGATE PM27 PROTEIN DISAPPEARS			
CC	FROM THE MORE MATURE MID-SHAFT REGIONS. MAY BE CONCENTRATED AT THE			
CC	MINERAL-NONMINERAL INTERFACE, RATHER THAN WITHIN THE SPICULE			
CC	MATRIX.			
CC	- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE			
CC	MICROME/PRIMARY MESENCHYME CELLS (PMCs) LINEAGE. PRODUCED			
CC	UNIFORMLY AND EXCLUSIVELY BY PMCS THROUGH THE EARLY PRISM STAGE			
CC	AND THIS SPECIFICITY IS FURTHER PESTICATED DURING SPLETOGENESIS			
CC	TO A SUBPOPULATION OF PMCS ASSOCIATED WITH THE GROWING TIPS OF THE			
CC	SPICULES.			
CC	- DEVELOPMENTAL STAGE: APPEARS AT THE MESENCHYME BLASTULA STAGE.			
CC	- DOMAIN: THE REPETITIVE DOMAIN MAY PROVIDE A CALCITRIE BINDING			
CC	MATRIX.			
CC	- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	EMBL: U01813; AA05689 1;			
CC	InterPro: IPR001304, Lectin_C.			
CC	PIfam: PF00059; lectin_c 1.			
CC	SMART: SM00034; CLECF 1.			
CC	PROSITE: PS00615; C-TYPE LECTIN 1; FALSE_NEG.			
CC	PROSITE: PS00041; C-TYPE LECTIN 2; 1.			
CC	Signal; Repeat.			

34	57	48.7	473	1	KICP_HUMAN	P10654 homo sapien
35	56.5	48.3	864	1	EIS_RAT	O99372 rattus norv
36	56	47.9	165	1	GRPI_ORYSA	P25074 oryza sativ
37	56	47.9	182	1	K2C3_BOVIN	P04261 bos taurus
38	56	47.9	238	1	EPF3_HUMAN	P52797 homo sapien
39	56	47.9	471	1	K1CN_HUMAN	P02533 homo sapien
40	56	47.9	502	1	K2C8_XENLA	P08776 xenopus lae
41	56	47.9	661	1	KMG_MOUSE	O08677 mus musculu
42	56	47.9	1250	1	TP3A_DROME	O79998 drosophila
43	56	47.9	2038	1	FSH_DROME	P13709 drosophila
44	55.5	47.4	341	1	ROA2_MOUSE	O08569 mus musculu
45	55.5	47.4	353	1	ROA2_HUMAN	P22626 homo sapien

FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 267 27 KDA PRIMARY MESOZYCHME-SPECIFIC
 FT DOMAIN 20 70 REPEAT-RICH REGION.
 FT DOMAIN 79 220 C-TYPE LECTIN.
 SO SEQUENCE 267 AA; 28589 MW; 2AC398C630DE5E07 CPGC64;

Query Match 60.7%; Score 71; DB 1; Length 267;
 Best Local Similarity 63.2%; Pred. No. 0.11;
 Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

CV 2 GVGAGFGNGVGPSTPGSG 20
 CH 39 GVGAGFGNGVGPSTPGSG 39

RESULT 2

MSA2_PLAF2

ID_MSA2_PLAF2

STANDARD

PRT; 347 AA

AC Q03646;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Merozoite surface antigen 2 precursor (MSA-2).

GN MSA2.

OC Plasmodium falciparum (isolate Nig32 / Nigeria)

OC Eukaryota, Alveolata, Apicomplexa, Haemosporidia, Plasmodium.

OX NCBI_TaxID=70150;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91156685; PubMed=2000183;

RA Smythe J A, Coffey P L, Day P P, Martin P R, Oduola A M J,

RA Kemp D J, Anders P F;

RT "Structural diversity in the Plasmodium falciparum merozoite surface

RT antigen 2";

RU Proc Natl Acad Sci U S A 88:1751-1755(1991).

CC -I- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE

CC -I- ERYTHROCYTE.

CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor

CC (Potential).

CC -I- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 58.1%; Score 68; DB 1; Length 347;
 Best Local Similarity 60.0%; Pred. No. 0.3;
 Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

CV 1 NGVAGFGNGVGPSTPGSG 20
 CH A4 NGVAGFGNGVGPSTPGSG 103

RESULT 3

GRP1_PETHY

ID_GRP1_PETHY

STANDARD

PRT; 384 AA

AC P09789;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 01-AUG-1992 (Rel. 23, Last annotation update)

DE Glycine-rich cell wall structural protein 1 precursor

GN GRP-1.

OS Petunia hybrida (petunia).

OC Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae, euasterids I; Solanales; Solanaceae; Petunia.

OX NCBI_TaxID=4102;

RN [1]

RP SEQUENCE FROM N.A.

RA Condit C M, Meagher R B;

RT "A gene encoding a novel glycine-rich structural protein of petunia";

RL Nature 323:178-181(1986)

CC -I- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).

CC -I- SUBCELLULAR LOCATION: Cell wall (potential).

CC -I- MISCELLANEOUS: THIS PROTEIN CONTAINS 678 GLYCINE RESIDUES.

CC -I- MISCELLANEOUS: 90% OF THE MATURE PROTEIN RESIDUES ARE CAPABLE OF

CC FORMING A BETA-SHEET COMPOSED OF 8 ANTI-PARALLEL STRANDS.

CC -I- MISCELLANEOUS: THE GLYCINE-RICH REGION IS COMPRISED OF TWO RELATED

CC FAMILIES OF REPEATS. F1 AND F2. EACH REPEAT CONTAINING ABOUT 40

CC AA.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 57.3%; Score 67; DB 1; Length 384;
 Best Local Similarity 68.4%; Pred. No. 0.42;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

CV 2 GVGAGFGNGVGPSTPGSG 20
 CH 360 GVGAGFGNGVGPSTPGSG 378

RESULT 4

FBOH_BOMMO

ID_FBOH_BOMMO

STANDARD

PRT; 5263 AA

AC P05750; Q26379; Q17220;

DT 01-NOV-1988 (Rel. 09, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Fibroin heavy chain precursor (Fib-H) (H-fibroin).

GN FIBH.

OS Bombyx mori (Silk moth).

OC Eukaryota, Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;

OC Ditrysia; Bombycoidea; Bombycidae; Bombyx.

NCBI_TaxID=7091;
 [1] SEQUENCE FROM N.A.
 RA ZHOU C.-C., Confalonieri F., Medina N., Zivanovic Y., Esnault C.,
 RA Yang T., Jaquet M., Janin J., Duquet M., Perasso P., Li Z.-G.,
 RA "Fine organization of Bombyx mori fibroin heavy chain gene.",
 RL Nucleic Acids Res. 28:2413-2419(2000).
 [2] SEQUENCE OF 1-168 FROM N.A.
 RA MEDLINE=80045039; PubMed=498286;
 RA Teujimoto Y., Suzuki Y.;
 RT "The DNA sequence of Bombyx mori fibroin gene including the 5'
 RT flanking, mRNA coding, putative intervening and fibroin protein coding
 RT regions";
 RL Cell 18:591-600(1979).
 [3] PARTIAL SEQUENCE FROM N.A.
 RA MEDLINE=79211211; PubMed=455439;
 RA Teujimoto Y., Suzuki Y.;
 RT "Structural analysis of the fibroin gene at the 5' end and its
 RT surrounding regions";
 RL Cell 16:425-436(1979).
 [4] PARTIAL SEQUENCE FROM N.A.
 RA STRAIN=Kinsu X Showa;
 RA MEDLINE=80994868; PubMed=3210244;
 RA Mita K., Ichimura S., Zama M., James T.C.;
 RA "Specific codon usage pattern and its implications on the secondary
 RA structure of silk fibroin mRNA";
 RL J. Mol. Biol. 203:917-925(1988).
 [5] PARTIAL SEQUENCE FROM N.A.
 RA MEDLINE=9435842; PubMed=7916056;
 RA Mita K., Ichimura S., James T.C.;
 RT "Highly repetitive structure and its organization of the silk fibroin
 RT gene";
 RL J. Mol. Evol. 38:583-592(1994).
 [6] SEQUENCE OF 517-523 FROM N.A. AND DISULFIDE BONDS
 RA STRAIN=J-139;
 RA MEDLINE=9296390; PubMed=10366732;
 RA Tanaka K., Yajiyama N., Ishikura K., Waga S., Yikuchi A., Ohtomo K.,
 RA Takagi T., Mizuno S.;
 RT "Determination of the site of disulfide linkage between heavy and
 RT light chains of silk fibroin produced by Bombyx mori.";
 RL Biochim Biophys. Acta 1432:92-103(1999).
 [7] FUNCTION: FORMS THE SILK FILAMENT: A STRONG, INEXTENSIBLE,
 [8] INSOLUBLE AND CHEMICALLY INERT FIBRE.
 [9] STRUNIT. FORMED OF TWO CHAINS: HEAVY AND LIGHT, THAT ARE LINKED BY
 [10] A DISULFIDE BOND. HEAVY-LIGHT CHAIN ASSEMBLY IS ESSENTIAL FOR THE
 [11] EFFICIENT INTRACELLULAR TRANSPORT AND SECRETION OF FIBROIN
 [12] TISSUE SPECIFICITY: PRODUCED EXCLUSIVELY IN THE POSTERIOR (PSG)
 [13] SECTION OF SILK GLANDS.
 [14] DOMAIN: COMPOSED OF ANTIPARALLEL BETA SHEETS. THE STRANDS OF THE
 [15] BETA SHEETS RUN PARALLEL TO THE FIBER AXIS. LONG STRETCHES OF SILK
 [16] FIBROIN ARE COMPOSED OF MICROCRYSTALLINE APAYS OF (GLY-SER-GLY-
 [17] ALA-GLY-ALA)-N INTERPEPTED BY REGIONS CONTAINING BULKIER RESIDUES
 [18] THE FIBER IS COMPOSED OF MICROCRYSTALLINE APAYS ALTERNATING WITH
 [19] AMORPHOUS REGIONS.
 [20] THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 [21] between the Swiss Institute of Bioinformatics and the EMBL outstation -
 [22] the European Bioinformatics Institute. There are no restrictions on its
 [23] use by non-profit institutions as long as its content is in no way
 [24] modified and this statement is not removed. Usage by and for commercial
 [25] entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 [26] or send an email to license@isb-sib.ch).
 [27] EMBL: AF226688; AAF76983.1;
 [28] EMBL: V00094; CAA23432.1;
 [29] EMBL: V00097; CAA23433.1;
 [30] EMBL: S74439; AAB31861.1;

DR EMBL: X13869; CAA32076.1;
 DR EMBL: M5378; AAA27839.1;
 DR EMBL: AB017362; BAA33147.1;
 DR PIR: S01844; S01844.
 KW Silk; Signal; Repeat.
 FT SIGNAL 1 21
 FT CHAIN 22 5263
 FT DOMAIN 149 5206
 FT DISULFID 5244 5244
 FT DISULFID 5260 5263
 FT CONFLICT 10 10
 FT SEQUENCE 5263 AA; 391586 MW; 8EE11D3A047440E CRC64;
 Query Match 56.4%; Score 66; DB 1; Length 5263;
 Best Local Similarity 68.4%; Pred No. 5.6;
 Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 2 GVGHFGNGVPGTGPQSG 20
 DB 848 GVGHFGNGVPGTGPQSG 866
 RESULT 5
 PHX5_MOUSE STANDARD; PRI: 672 AA.
 ID PHX5_MOUSE
 AC P08399;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Per-hexamer repeat protein 5.
 GN PHX5 OR PER.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 FN [1]
 PP SEQUENCE FROM N.A.
 PX MEDLINE=86014364; PubMed=2413365;
 PX Shin H.S., Bargiello T.A., Clark B.T., Jackson F.R., Young M.W.;
 RT "An unusual coding sequence from a Drosophila clock gene is conserved
 RT in vertebrates";
 RL Nature 317:445-448(1985).
 [1] SIMILARITY: TO OTHER G-T STRETCH CONTAINING PROTEINS.
 [2] THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 [3] between the Swiss Institute of Bioinformatics and the EMBL outstation -
 [4] the European Bioinformatics Institute. There are no restrictions on its
 [5] use by non-profit institutions as long as its content is in no way
 [6] modified and this statement is not removed. Usage by and for commercial
 [7] entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 [8] or send an email to license@isb-sib.ch).
 [9] EMBL: M12039; AAA88320.1;
 DR EMBL: X02966; CAA26710.1; ALT_INIT.
 DR PIR: A24403; UMMS.
 DR MGD; MGI:104521; Phx5.
 DR InterPro; IPR000561; EGF-like.
 DR SMART; SM00181; EGF; 1.
 DR POSSITE; PSM1166; EGF_2; UNKNOWN_1.
 KM Repeat.
 FT DOMAIN 59 672
 FT SEQUENCE 672 AA; 57924 MW; E85BF428CF424C0B CRC64;
 Query Match 54.7%; Score 64; DB 1; Length 672;
 Best Local Similarity 63.2%; Pred No. 1.5;
 Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 2 GVGHFGNGVPGTGPQSG 20
 DB 326 GTGFGSGSGTGTGTGSG 344
 RESULT 6

HOLD1
ID HOLD1 STANDARD; PRT; 104 AA.
AC Q25055;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Holotricin 3 precursor.
OS Holotricin diomphalia.
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Scarabaeiformia; Scarabaeidae; Melolonthinae; Holotricina
NCBI_TaxID=33394;
RN [1]
PP SEQUENCE FROM N.A. AND SEQUENCE OF 21 40.
PC TISSUE=Larval hemolymph;
RX MEDLINE=96073722; PubMed=8535393;
RA Lee S.Y., Moon H.J., Kurata S., Natori S., Lee B.L.;
RT "Purification and cDNA cloning of an antifungal protein from the
hemolymph of Holotricia diomphalia larvae."
RL Biol. Pharm. Bull. 18:1049-1052(1995).
CT FUNCTION: HOLOTRICIN 3 HAS ANTIFUNGAL ACTIVITY AGAINST
CC C ALPHACANS.
CC -1 SIMILARITY: TO TENECIN 3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: D13744; BAA02889.1;
KW Insect immunity; Antidiotic; Hemolymph; Fungicide; Signal; Repeat.
FT SIGNAL 1 20 HOLOTICIN 3.
FT PEPTIDE 21 104 18 X 4 AA APPROXIMATE TANDEM REPEATS OF
FT DOMAIN 27 98 H-G-G-G.
FT REPEAT 27 30 1.
FT REPEAT 31 34 2.
FT REPEAT 35 38 3.
FT REPEAT 39 42 4.
FT REPEAT 43 46 5.
FT REPEAT 47 50 6.
FT REPEAT 51 54 7.
FT REPEAT 55 58 8.
FT REPEAT 59 62 9.
FT REPEAT 63 66 10.
FT REPEAT 67 70 11.
FT REPEAT 71 74 12.
FT REPEAT 75 78 13.
FT REPEAT 79 82 14.
FT REPEAT 83 86 15.
FT REPEAT 87 90 16.
FT REPEAT 91 94 17.
FT REPEAT 95 98 18.
SQ SEQUENCE 104 AA; 9026 MW; 2799D681BFDC725 CRC64;
Query Match 53.8%; Score 63; DB 1; Length 104;
Best Local Similarity 60.0%; Pred. No. 0.38;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
DY 1 NCVRHGFQNGVGFPTGPGSG 23
Dh 35 HGGGHGNGGQGGHGHGHPDGG 54

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Secretory granule proteoglycan core protein precursor (platelet
DE proteoglycan core protein) (P.PG) (Hematopoietic proteoglycan core
DE protein) (Serglycin).
GN PPG1 OR PPG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NCBI_TaxID=9606;
RN [1]
PP SEQUENCE FROM N.A.
RX MEDLINE=90016819; PubMed=2798108;
RA Steller C.M., Saunders G.F.;
RT "Nucleotide sequence of a cDNA encoding a hemopoietic proteoglycan
RT core protein."
RL Nucleic Acids Res. 17:7523-7523(1989).
RN [2]
PP SEQUENCE FROM N.A.
RX MEDLINE=8823411; PubMed=2815370;
RA Stevens R.L., Avraham S., Gartner M.C., Bruns G.A.P., Austen K.F.,
RA Weiss J.H.;
RT "Isolation and characterization of a cDNA that encodes the peptide
RT core of the secretory granule proteoglycan of human promyelocytic
RT leukemia HL-60 cells."
RL J. Biol. Chem. 263:7287-7291(1988).
RN [3]
PP SEQUENCE FROM N.A.
RX MEDLINE=90202841; PubMed=2180935;
RA Nicodemus C.F., Avraham S., Austen K.F., Fudys S., Jablonski J.,
RA Stevens R.L.;
RT "Characterization of the human gene that encodes the peptide core of
RT secretory granule proteoglycan in promyelocytic leukemia HL-60 cells
RT and analysis of the translated product."
RL J. Biol. Chem. 265:5889-5896(1990).
RN [4]
PP SEQUENCE FROM N.A.
PC TISSUE=Lung;
RA Strauberg R.;
RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
PP SEQUENCE OF 28-158, AND SEQUENCE OF 34-158 FROM N.A.
RX MEDLINE=88296856; PubMed=3402609;
RA Alliel P.M., Perin J.-P., Maillet P., Bonnet F., Rosa J.-P.,
RA Jolles P.;
RT "Complete amino acid sequence of a human platelet proteoglycan."
RL FEBS Lett. 236:123-126(1988).
RN [6]
PP SEQUENCE OF 28-93.
RX MEDLINE=89104992; PubMed=3214420;
RA Perin J.-P., Bonnet F., Maillet P., Jolles P.;
RT "Characterization and N-terminal sequence of human platelet
RT proteoglycan."
RL Biochem. J. 255:1007-1013(1988).
CC -1 FUNCTION: MAY NEUTRALIZE HYDROLYTIC ENZYMES.
CC -1 INDUCTION: BY EBV.
CC -1 SIMILARITY: TO CORRESPONDING PROTEINS IN RAT AND MOUSE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M33651; AAA60322.1;
CC EMBL: M33649; AAA60322.1; JOINED.
CC EMBL: M33650; AAA60322.1; JOINED.
DR EMBL: X17042; CAA14900.1;
DR EMBL: X12765; CAA11255.1;
DR EMBL: BC015516; AAH15516.1;
DR EMBL: U03223; AAA60179.1;
DR PIR: A28058; A28058.

DR PIR: A35183; A35183.
 DR PIR: S01126; S01126.
 DR PIR: S01819; S01819.
 DR PIR: S00610; S00610.
 DR Genew: HGNC:9361; PRG1.
 DR MIM: 177040; .
 KW Glycoprotein; Proteoglycan; Repeat; Signal
 FT SIGNAL 1 27
 FT CHAIN 28 158
 FT DOMAIN 94 111
 FT CARBOHYD 94 94
 FT CARBOHYD 96 96
 FT DISULFID 40 49
 FT CONFLICT 139 139
 FT SEQUENCE 158 AA; 17624 MW; 1275B7F39FF91476 CRC64;

Query Match 52.1%; Score 61; DB 1; Length 158;
 Best Local Similarity 64.7%; Pred. No. 0.93;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 GHGFGNGVGTGPGSG 20
 Db 95 GSGFGSGSGSGSGSG 111

RESULT 8

GRP_ARATH
 ID GRP_ARATH STANDARD; PRT; 338 AA.

AC P27483;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DE Glycine-rich cell wall structural protein precursor.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OS eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 CX NCB1_TaxId=1072;
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=92003708, PubMed=1912511;
 PA Quigley F., Willitt M.L., Macho P.;
 RT "Nucleotide sequence and expression of a novel glycine-rich protein
 gene from Arabidopsis thaliana".
 RL Plant Mol. Biol. 17:949-952(1991).
 CC -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Cell wall (Potential).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X58838; CAA41249.1;
 DR PIR: S17732; KMMO.
 KW Cell wall; Structural protein; Repeat; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 338
 FT DOMAIN 21 338
 FT SEQUENCE 338 AA; 23891 MW; 046A5E8C1AEB89EB CRC64;

Query Match 52.1%; Score 61; DB 1; Length 338;
 Best Local Similarity 63.2%; Pred. No. 1.8;
 Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 GVCHFGNGVGTGPGSG 20

Db 295 GVGGFGGSGGGGFGGAG 313

RESULT 9

PER_DROTE
 ID PER_DROTE STANDARD; PRT; 88 AA.

AC Q26287;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Period circadian protein (Fragment).
 GN PER.
 OS Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OS Insecta; Preygoia; Neoptera; Endopterygota; Diptera; Brachycera;
 OS Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CX NCB1_TaxId=7243;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93140158; PubMed=1487825;
 RA Peixoto A.A., Costa R., Wheeler D.A., Hall J.C., Kyriacou G.P.;
 RT "Evolution of the theonine-glycine repeat region of the period gene
 in the melanogaster species subgroup of Drosophila".
 RL J. Mol. Evol. 35:411-419(1992).
 CC -1- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
 CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN
 CC PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
 CC RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE
 CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
 CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
 CC PHYTIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
 CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES
 CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
 CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
 CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
 CC TRANSCRIPTIONAL LOOP DOES NOT APPEAR TO BIND DNA, SUGGESTING
 CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM). THE COMPLEX THEN
 CC TRANSLOCATES INTO THE NUCLEUS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
 CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
 CC TRANSLLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
 CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
 CC -1- DOMAIN: THE PIR OF GUY-THP IS IMPLICATED IN THE MAINTENANCE OF
 CC CIRCADIAN PERIOD AT DIFFERENT TEMPERATURES. DELETION OF THE REPEAT
 CC LEADS TO A SHORTENING OF THE COURSE SONG CYCLE PERIOD, AND THUS
 CC COULD BE IMPORTANT FOR DETERMINING FEATURES OF SPECIES-SPECIFIC
 CC MATING BEHAVIOR (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE
 CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
 CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
 CC PER-TIM (BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: S53300; AAB25030.2;
 DR FlyBase: FBgn0013030; Dmel/per.
 KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
 FT NON_TER 1 1
 FT DOMAIN 33 61
 FT NON_TER 88 88
 FT SEQUENCE 88 AA; 7961 MW; 614FC71773AF1B33 CRC64;

Query Match 51.3%; Score 60; DB 1; Length 88;
 Best Local Similarity 57.9%; Pred. No. 0.72;
 Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;


```

DT 16-OCT-2001 (Rel.. 40, last annotation update)
DE Keratin, type I cytoskeletal 47 kDa.
GN XK81B2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxId=8355;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=87057649; PubMed=2430981,
RA Myltant S., Minkles J.A., Sargent T.D., David I.R.;
RT "Stage-specific keratins in Xenopus laevis embryos and tadpoles: the
RT XK81 gene family."
RL J. Cell Biol. 103:1957-1965(1996).
CC 1- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC 1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFILLLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
CC 1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-eb.ch/announce/
CC or send an email to license@ib-eb.ch).
CC
DR EMBL; M8155; -; NOT ANNOTATED COS.
DR EMBL; X04807; CAA28498.1; -
DR EMBL; X12730; CAA31223.1; -
DR EMBL; X05865; CAA29293.1; -
DR PIR; A25438; A25438.
DR InterPro; IPR001664; IF.
DR Pfam; PF000398; filament_1
DR Pfam; PF000397; Keratin_1.
DR PRINTS; PR01248; TYPE1KERATIN.
DR PROSITE; PS00226; IF_1.
KM Intermediate filament; Coiled coil; Keratin.
FT DOMAIN 1 81 HEAD
FT DOMAIN 82 393 ROD
FT DOMAIN 394 419 TAIL.
FT DOMAIN 82 117 COIL 1A.
FT DOMAIN 118 139 LINER 1.
FT DOMAIN 140 231 COIL 1B.
FT DOMAIN 232 254 LINER 12.
FT DOMAIN 255 393 COIL 2.
SQ SEQUENCE 419 AA; 45652 MW; BA16F03BEF51D415A CPr664;

Query Match 51.3%; Score 60; DB 1; Length 419;
Best Local Similarity 63.2%; Pred. No. 2.9;
Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GVSHGFNGCVGPGTGPSSG 20
DB 21 GYGAGFGGSGSGAGFGGSG 39

Search completed: May 6, 2003, 14:49:13
Job time: 13 secs

```

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:47:42 ; Search time 15 Seconds
(without alignments)
128.179 Million cell updates/sec

Title: US-09-926-234-1

Perfect score: 117

Sequence: 1 NGVGHGFGNGVPGTGPSSG 20

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR 73:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	72	61.5	532 2 T35119	probable aminotran
2	70	59.8	229 2 T27840	hypothetical prote
3	70	59.8	259 2 T17889	glycine-rich prote
4	68	58.1	137 2 T04930	glycine-rich cell
5	68	58.1	185 2 JC4085	glycine-rich cutic
6	68	58.1	347 2 B39112	metazolate 45K surf
7	68	58.1	617 2 T49444	lustrin A related
8	67	57.3	384 1 A26099	glycine-rich cell
9	66	56.4	215 2 S54203	keratin 6 - bovine
10	65	55.6	396 2 T49109	glycine-rich prote
11	65	55.6	1226 2 T24045	hypothetical prote
12	64	54.7	100 2 T17558	period clock prote
13	64	54.7	713 1 UNMS	holocrine 3 precu
14	63	53.8	104 2 JC4190	molluscan shell m
15	63	53.8	140 2 UC7211	probable cell wall
16	63	53.8	205 2 T07959	eyelid - fruit fly
17	63	53.8	2715 2 T13049	fibroin - silkworm
18	62	53.0	108 1 S01844	glycine-rich prote
19	62	53.0	210 2 J01060	hypothetical prote
20	62	53.0	504 2 T13475	ferredoxin - Therm
21	61	52.1	95 2 H72271	glycine-rich prote
22	61	52.1	107 2 B83366	hypothetical prote
23	61	52.1	127 2 B82734	hypothetical prote
24	61	52.1	129 2 T17530	glycine-rich prote
25	61	52.1	158 2 A28058	secretory granule
26	61	52.1	173 2 J01064	glycine-rich prote
27	61	52.1	220 2 T14441	glycine-rich prote
28	61	52.1	338 1 KMTU	glycine-rich cell
29	61	52.1	408 2 S57483	glycin-rich protei

30	61	52.1	457 2 S39079	puff C-8 protein -
31	61	52.1	629 2 S42629	keratin K3 - rabb
32	61	52.1	764 2 H71607	hypothetical prote
33	61	52.1	1473 2 T13855	suppressor of babl
34	60	51.3	111 2 T14306	glycine-rich prote
35	60	51.3	170 2 JC2213	hypothetical 14.7K
36	60	51.3	174 2 S00273	period clock prote
37	60	51.3	201 2 T00799	hypothetical prote
38	60	51.3	251 2 D96010	hypothetical prote
39	60	51.3	255 2 B84777	hypothetical prote
40	60	51.3	291 1 S31415	glycine-rich prote
41	60	51.3	314 2 F75504	probable plasm. ty
42	60	51.3	419 2 A25438	keratin, type I cy
43	60	51.3	571 1 UMFF	period clock prote
44	60	51.3	1122 2 B26427	period clock prote
45	60	51.3	1127 2 A25018	circadian rhythm p

ALIGNMENTS

RESULT 1
T35119
Probable aminotransferase - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C/Accession: T35119
R/Seeger, K.D.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, March 1998
A/Reference number: Z21568
A/Accession: T35119
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-532 <SEE>
A/Cross-references: EMBL:AL022268; PIDN:CA18336.1; GSPDR:GN00070; SGOEDB:SC4H2.2
A/Experimental source: strain A3(2)
A/Genetic:
A/Gene: SGOEDB:SC4H2.20

Query Match 61.5%; Score 72; DB 2; Length 532;
Best Local Similarity 68.4%; Pred. No. 0.23;
Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 GVGHGFGNGVPGTGPSSG 20
DB 84 GSGTGPRTGTPGTGPGTG 102

RESULT 2

T27840
Hypothetical protein ZK39.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T27840
R/Kershaw, J.
Submitted to the EMBL Data Library, November 1996
A/Reference number: Z20428
A/Accession: T27840
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-329 <WIL>
A/Cross-references: EMBL:Z82093; PIDN:CAB05018.1; GSPDR:GN00019; CEST:ZK39.2
A/Experimental source: clone ZK39
C/Genetics:
A/Gene: CEST:ZK39.2
A/Map position: 1
A/Introns: 28/1; 55/3; 111/1

Query Match 59.8%; Score 70; DB 2; Length 229;
Best Local Similarity 65.0%; Pred. No. 0.18;
Matches 13; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVPGTGPSSG 20

Db 30 NGCHGHGNGNGNGNGNG 49

RESULT 3

glycine-rich protein a386r - Chlorella virus PBCV-1
 C:Species: Chlorella virus PBCV-1
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 R:Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL Data Library, May 1999
 A:Accession: T17889
 A:Status: preliminary; transferred from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-259 <GRA>
 A:Cross-references: EMBL:U42580, NID:94028896, PID:AA096754.1
 A:Experimental source: specific host Chlorella strain NC64A
 C:Genetics:
 A:Note: a386r

Query Match 59.8%; Score 70; DB 2; Length 259;
 Best Local Similarity 70.6%; Pred. No. 0.2;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RESULT 4

glycine-rich cell wall protein homolog T9A21.130 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 18-Feb-2000
 C:Accession: T04930
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, February 1999
 A:Reference number: Z15390
 A:Accession: T04930
 A:Molecule type: DNA
 A:Residues: 1-137 <BEV>
 A:Cross-references: EMBL:AL021713
 A:Experimental source: cultivar Columbia; BAC clone T9A21
 C:Genetics:
 A:Map position: 4
 A:Note: T9A21.130
 C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match 58.1%; Score 68; DB 2; Length 137;
 Best Local Similarity 68.4%; Pred. No. 0.19;
 Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 5

glycine-rich cuticular protein precursor - yellow mealworm
 C:Species: Tenebrio molitor (yellow mealworm)
 C:Date: 12-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 24-Nov-1999
 C:Accession: J04085
 R:Mathelin, J.; Bounin, H.; Quennedey, B.; Courrent, A.; Delachambre, J.
 Gene 156, 259-264, 1995
 A:Title: Identification, sequence and mRNA expression pattern during metamorphosis of a
 A:Reference number: J04085; MUID:95278754; PMID:7758965
 A:Accession: J04085
 A:Molecule type: mRNA
 A:Residues: 1-185 <MAT>
 A:Cross-references: GB:J36203, NID:9537944; PID:9537945

A:Experimental source: epidermal cells
 C:Superfamily: glycine-rich cell wall structural protein 1
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-185/Product: glycine-rich cuticular protein #status predicted <MA2>
 F:28-34, 81-87, 145-151/Region: 7-residue repeats
 F:69-73, 75-79/Region: 5-residue repeats

Query Match 58.1%; Score 68; DB 2; Length 185;
 Best Local Similarity 68.4%; Pred. No. 0.26;
 Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 6

merozoite 45K surface antigen precursor - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 17-Nov-2000
 C:Accession: B39112
 R:Smyle, J.A.; Coppel, R.L.; Day, K.P.; Martin, P.K.; Chulaj, A.M.U.; Kemp, D.J.
 Proc. Natl. Acad. Sci. U.S.A. 88, 1751-1755, 1991
 A:Title: Structural diversity in the Plasmodium falciparum merozoite surface anti
 A:Reference number: A39112, MUID:9156685; PMID:2000383
 A:Accession: B39112
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-347 <SMY>
 A:Cross-references: GB:M59767
 C:Superfamily: Epstein-Barr virus nuclear antigen
 C:Keywords: surface antigen

Query Match 58.1%; Score 68; DB 2; Length 347;
 Best Local Similarity 60.0%; Pred. No. 0.46;
 Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

RESULT 7

luciferin A related protein [imported] - Neurospora crassa
 N:Alternate names: protein B17C10.250
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49444
 R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, K.; Nyat
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49444
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-617 <SCH>
 A:Cross-references: EMBL:AL355926; GSPDB:GN00116; NCSP:B17C10.250
 A:Experimental source: BAC clone B17C10; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B17C10.250
 A:Map position: 6

Query Match 58.1%; Score 68; DB 2; Length 617;
 Best Local Similarity 65.0%; Pred. No. 0.78;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

RESULT 8

A26099
glycine-rich cell wall structural protein - garden petunia
C/Species: Petunia x hybrida (garden petunia)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: A26099
R/Condit, C.M.; Meagher, R.B.
Nature 323, 178-181, 1986
A/Title: A gene encoding a novel glycine-rich structural protein of petunia.
A/Reference number: A26099
A/Accession: A26099
A/Molecule type: mRNA
A/Residues: 1384 <CON>
C/Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 57.4% Score 67; DB 1; Length 384;
Best Local Similarity 68.4% Pred. No. 0.66;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Dh 360 GVGAGAGHGVGVGSSGSG 378

RESULT 9
S54203
Keratin 6 - bovine (fragment)
C/Species: Bos primigenius taurus (cattle)
C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C/Accession: S54203
R/Navarro, M.; Segreles, C.; Jorcano, J.L.
submitted to the EMBL Data Library, February 1995
A/Description: Identification of a new bovine keratin 6 gene.
A/Reference number: S54201
A/Accession: S54203
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1215 <NAV>
A/Cross-reference: EMBL:248498, NID:G703930, PIRN:CAA8498.1; PIR:G791811
C/Genetics:
A/Intons: 188/3
C/Superfamily: cytokeleletal keratin

Query Match 56.4% Score 66; DB 2; Length 215;
Best Local Similarity 63.2% Pred. No. 0.51;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 GVGAGFNGVGPPTGPGSG 20
|||:|||||:|||||:
Db 87 GVGAGFNGVGPPTGPGSG 105

RESULT 10
T49109
glycine-rich protein - Arabidopsis thaliana
N/Alternate names: protein AT4G22020
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
C/Accession: T49109
R/Bevan, M.; Medler, H.; Manbut, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; M.
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25016
A/Accession: T49109
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1396 <BEV>
A/Cross-references: EMBL:AL022140, GSPDB:GNC0662, ATSP:AT4G22020
A/Experimental source: cultivar Columbia; BAC clone PIN20
C/Genetics:
A/Gene: ATSP:AT4G22020
A/Map position: 4
C/Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 55.6% Score 65; DB 2; Length 396;

Best Local Similarity 60.0% Pred. No. 1.2;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 NGVGHFGNGVGPPTGPGSG 20
|||:|||||:|||||:
Db 302 NGVGHFGNGVGPPTGPGSG 321

RESULT 11
T24045
hypothetical protein R08B4.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C/Accession: T24045
R/White, S.
submitted to the EMBL Data Library, November 1995
A/Reference number: Z19834
A/Accession: T24045
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 11226 <WIL>
A/Cross-references: EMBL:Z68008, PIRN:CAA92000.1; GSPDB:GN0002A; CESP:R08B4.1
A/Experimental source: clone R08B4
C/Genetics:
A/Gene: CESP:R08B4.1
A/Map position: X
A/Intons: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 810/3; 877/1; 9

Query Match 55.6% Score 65; DB 2; Length 1226;
Best Local Similarity 60.0% Pred. No. 3.3;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 NGVGHFGNGVGPPTGPGSG 20
|||:|||||:|||||:
Db 921 NGVGHFGNGVGPPTGPGSG 940

RESULT 12
T17558
glycine-rich protein a68L - Chlorella virus PCV-1
C/Species: Chlorella virus PCV-1
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T17558
R/Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A/Reference number: Z18806
A/Accession: T17558
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1100 <GRA>
A/Cross-references: EMBL:U42580; NID:G4028896; PIRN:AA096436.1
A/Experimental source: specific host Chlorella strain NC64A
C/Genetics:
A/Gene: a68L

Query Match 54.7% Score 64; DB 2; Length 100;
Best Local Similarity 63.2% Pred. No. 0.43;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 GVGAGFNGVGPPTGPGSG 20
|||:|||||:|||||:
Db 38 GVGAGFNGVGPPTGPGSG 56

RESULT 13
UMMS
period clock protein - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jun-2000
C/Accession: A24403
R/Shin, H.S.; Bargiello, T.A.; Clark, B.T.; Jackson, F.R.; Young, M.W.
Nature 337, 445-448, 1985
A/Title: An unusual coding sequence from a Drosophila clock gene is conserved in

A:Reference number: A24403; MUID:86014384; PMID:2413365
 A:Accession: A24403

A:Molecule type: DNA

A:Residues: 1-713 <SHI>

A:Cross-references: GB:X02966; GB:M12039, NID:G55125; PIGN:CAA26710.1; PID:G1334150

C:Comment: Mutations within the per locus of the fruit fly affect a variety of natural

logous locus with multiple tandem repeats of nucleic acid hexamers (ACNCGN, TCAGGC) that

C:Comment: The serine residues of the S-G repeats found in certain proteoglycans are at

C:Superfamily: period clock proteins; EGF homology

C:Keywords: circadian rhythm; tandem repeat

F:41-77/Domain: EGF homology <EGF>

Query Match

Best Local Similarity 54.7%; Score 64; DB 1; Length 713;
 Matches 12; Conservativity 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVGHGFGNGVGPGTGPGSG 20
 DB 367 GTGTGSGGTGTGTGTGSG 385

RESULT 14

JC4190

holotricin 3 precursor - Holotrichia diomphalia

N:Alternate names: antifungal protein

C:Species: Holotrichia diomphalia

C>Date: 04-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C:Accession: JC4190

R:Lee, S.Y.; Moon, H.J.; Kurata, S.; Natori, S.; Lee, B.L.

Biol. Pharm. Bull. 18, 1049-1052, 1995

A:Title: Purification and cDNA cloning of an antifungal protein from the hemolymph of Hd

A:Reference number: JC4190; MUID:96073722; PMID:8535393

A:Accession: JC4190

A:Molecule type: mRNA

A:Residues: 1-104 <LEES>

A:Cross-references: DDBJ:D13744; NID:G1088433; PIDN:BA402889.1; PID:d1003394; PID:G17861

C:Comment: This protein is a Gly- and His-rich protein and a constitutive protein of Jar

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-104/Product: holotricin 3 #status predicted <MAT>

Query Match

Best Local Similarity 53.8%; Score 63; DB 2; Length 104;
 Matches 12; Conservativity 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVGHGFGNGVGPGTGPGSG 20
 DB 35 HRRHHNNRQGGHGHGPGSG 54

RESULT 15

JC7211

molluscan shell matrix protein N14 - Pinctada maxima

C:Species: Pinctada maxima

C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000

C:Accession: JC7211

R:Kono, M.; Hayashi, N.; Samata, T.

Biochem. Biophys. Res. Commun. 269, 213-218, 2000

A:Title: Molecular mechanism of the nacreous layer formation in Pinctada maxima.

A:Reference number: JC7210; MUID:20160475; PMID:10694502

A:Accession: JC7211

A:Molecule type: mRNA

A:Residues: 1-140 <KON>

A:Cross-references: DDBJ:AB032613

C:Comment: This protein is rich in Gly, Tyr and Asn residues. It involves in the matrix

C:Keywords: matrix protein

Query Match

Best Local Similarity 53.8%; Score 63; DB 2; Length 140;
 Matches 12; Conservativity 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVGPGTGPGSG 20
 DB 111 GVGHGFGNGVGPGTGPGSG 20

DB 92 NGNGNGNGGYGNGNGNG 111

Search completed: May 6, 2003, 14:50:11
 Job time: 17 secs

GenCore version 5.1.4.D5.4578
Copyright (c) 1993 2003 Comugen Ltd.

CM protein - protein search, using sw model

Run on: May 6, 2003, 14:49:02 ; Search time 17 Seconds
(without alignments)

101,514 Million cell updates/sec

Title: US-09-926-234-1

Sequence: 117

Sequence: 1 NGVGHGFGNGVGPCTGPGSG 20

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328255 seqs, 8628685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database:

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	72	61.5	287	US-10-176-847-66
2	61	52.1	59	US-09-796-692-1354
3	61	52.1	59	US-09-796-692-1354
4	61	52.1	59	US-10-040-862-1354
5	61	52.1	59	US-10-040-862-1354
6	61	52.1	125	US-09-796-692-1354
7	61	52.1	125	US-09-796-692-1354
8	61	52.1	125	US-10-040-862-1354
9	61	52.1	125	US-10-040-862-1354
10	61	52.1	148	US-09-796-692-1354
11	61	52.1	148	US-10-040-862-1354
12	61	52.1	124	US-09-881-752A-118
13	61	52.1	156	US-09-796-692-1357
14	61	52.1	156	US-10-040-862-1357
15	61	52.1	162	US-09-925-306-1362
16	61	52.1	264	US-09-949-842-25
17	61	52.1	692	US-09-899-471-8
18	61	52.1	705	US-09-874-503-14
19	61	52.1	705	US-10-000-157-14

20	61	52.1	705	US-10-063-547-162	Sequence 162, App
21	61	52.1	705	US-09-816-744-14	Sequence 14, Appl
22	61	52.1	705	US-09-747-259-14	Sequence 14, Appl
23	61	52.1	705	US-10-174-590-598	Sequence 598, App
24	61	52.1	705	US-10-176-758-598	Sequence 598, App
25	61	52.1	705	US-10-063-616-162	Sequence 162, App
26	61	52.1	705	US-10-175-737-598	Sequence 598, App
27	61	52.1	705	US-10-063-502-162	Sequence 162, App
28	61	52.1	705	US-10-173-706-598	Sequence 598, App
29	61	52.1	705	US-10-175-752-598	Sequence 598, App
30	61	52.1	705	US-10-175-752-598	Sequence 598, App
31	61	52.1	705	US-10-176-482-598	Sequence 598, App
32	61	52.1	705	US-10-176-757-598	Sequence 598, App
33	61	52.1	705	US-10-175-737-598	Sequence 598, App
34	61	52.1	705	US-10-180-552-598	Sequence 598, App
35	61	52.1	705	US-10-180-557-598	Sequence 598, App
36	61	52.1	705	US-10-173-700-598	Sequence 598, App
37	61	52.1	705	US-10-174-572-598	Sequence 598, App
38	61	52.1	705	US-10-174-579-598	Sequence 598, App
39	61	52.1	705	US-10-174-582-598	Sequence 598, App
40	61	52.1	705	US-10-174-588-598	Sequence 598, App
41	61	52.1	705	US-10-175-739-598	Sequence 598, App
42	61	52.1	705	US-10-175-740-598	Sequence 598, App
43	61	52.1	705	US-10-175-743-598	Sequence 598, App
44	61	52.1	705	US-10-176-488-598	Sequence 598, App
45	61	52.1	705	US-10-176-492-598	Sequence 598, App

ALIGNMENTS

RESULT 1
US-10-176-847-66
Sequence 66, Application US/10176847
Publication No. US20030068636A1
GENERAL INFORMATION:
APPLICANT: Velby, Peter Ole
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF 1
FILE REFERENCE: MRI-039
CURRENT APPLICATION NUMBER: US/10/176,847
CURRENT FILING DATE: 2002-06-21
NUMBER OF SEQ ID NOS: 112
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 66
LENGTH: 287
TYPE: PRT
ORGANISM: Homo sapiens
US-10-176-847-66

Query Match: 61.5% Score 72; DB 9; Length 287;
Best Local Similarity 63.6%; Pred No. 0.18;
Matches 14; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 1 NGVGHGFGNGVGPCTG--PQSG 20
Db 144 SGAGHSGSGSGPGCGSVPGSG 165

RESULT 2
US-09-796-692-1354
Sequence 1354, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Aigate, Paul A.
TITLE OF INVENTION: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND
FILE REFERENCE: 2077,001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01

```

1 PRIOR APPLICATION NUMBER: 60/186,126
2 PRIOR FILING DATE: 2000-03-01
3 PRIOR APPLICATION NUMBER: 60/190,479
4 PRIOR FILING DATE: 2000-03-17
5 PRIOR APPLICATION NUMBER: 60/200,545
6 PRIOR FILING DATE: 2000-04-27
7 PRIOR APPLICATION NUMBER: 60/200,303
8 PRIOR FILING DATE: 2000-04-28
9 PRIOR APPLICATION NUMBER: 60/200,779
10 PRIOR FILING DATE: 2000-04-28
11 PRIOR APPLICATION NUMBER: 60/200,999
12 PRIOR FILING DATE: 2000-05-01
13 PRIOR APPLICATION NUMBER: 60/202,084
14 PRIOR FILING DATE: 2000-05-04
15 PRIOR APPLICATION NUMBER: 60/206,201
16 PRIOR FILING DATE: 2000-05-22
17 PRIOR APPLICATION NUMBER: 60/218,950
18 PRIOR FILING DATE: 2000-07-14
19 PRIOR APPLICATION NUMBER: 60/222,903
20 PRIOR FILING DATE: 2000-08-03
21 PRIOR APPLICATION NUMBER: 60/223,416
22 PRIOR FILING DATE: 2000-08-04
23 PRIOR APPLICATION NUMBER: 60/223,378
24 PRIOR FILING DATE: 2000-08-07
25 NUMBER OF SEQ ID NOS: 9597
26 SOFTWARE: FastSeq for Windows Version 3.0
27 SEQ ID NO: 1354
28 LENGTH: 59
29 TYPE: PRT
30 ORGANISM: Homo sapiens
31 US-00-736-692-1354

```

```

Query Match      52.1%   Score 61, DB 9,   Length 59
Best Local Similarity 64.7%   Pred. No. 0.78;
Matches 11; Conservative 2; Mismatches 4; Indels
QY      4  GHSFGNGVSPGTGPGSG 20
      | | | | | | | | | |
Db       33  GSGPFGSGSGSGSGSG 49
      | | | | | | | | | |

```

RESULT 3
US-09-796-692-2320
Sequence 2320. Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077 001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903

```

PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FASTSEQ for Windows Version 3.0.0
SEQ ID NO: 3120
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-692.2320

```

```

Query Match      52.1%   Score 61, DB 9; Length 59;
Best Local Similarity 64.7%   Pred. No. 0.78;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      4 GAGCGNGCGCTGCGG 20
      | | | | | | | | | |
Db      33 GCGCGGCGGCGGCGG 49

```

```

RESULT 4
US-10-040-862-1354
: Sequence 1354, Application US/10040862
: Publication No. US20030078396A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: APPLICANT: Retter, Marc
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and
: FILE REFERENCE: 014058-013520US
: CURRENT APPLICATION NUMBER: US/10/040,862
: CURRENT FILING DATE: 2001-11-06
: PRIOR APPLICATION NUMBER: US 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: US 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: US 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: US 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: US 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: US 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: US 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: US 60/223,378
: PRIOR FILING DATE: 2000-08-07
: PRIOR APPLICATION NUMBER: US 09/796,692
: PRIOR FILING DATE: 2001-03-01
: NUMBER OF SEQ ID NOS: 10467
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1354
: LENGTH: 59
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-040-862-1354

Query Match      52.1%   Score 61;   DB 9;   Length 59;
Best Local Similarity  64.7%;   Pred. No. 0.78;

```

```

; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04

```

```
ORGANISM: Homo sapiens
US-10-040-862-1301
Query Match
Best Local Similarity 52.1% Score 61; DB 9; Length 125;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2292
LENGTH: 125
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-692-2292
```

```
Query Match
Best Local Similarity 52.1% Score 61; DB 9; Length 125;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 4 GHGFGNGVPGTGPSCG 20
Db 62 GSGFGSGSGSGSGSGSG 78
```

```
RESULT 8
US-10-040-862-1301
Sequence 1301, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Aigate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-01352005
CURRENT APPLICATION NUMBER: US/10/040,862
PRIOR FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1301
LENGTH: 125
TYPE: PRT
```

```
ORGANISM: Homo sapiens
US-10-040-862-1301
```

```
Query Match
Best Local Similarity 52.1% Score 61; DB 9; Length 125;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 4 GHGFGNGVPGTGPSCG 20
Db 62 GSGFGSGSGSGSGSGSG 78
```

```
RESULT 9
US-10-040-862-2292
Sequence 2292, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Aigate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and
FILE REFERENCE: 014058-01352005
CURRENT APPLICATION NUMBER: US/10/040,862
PRIOR FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2292
LENGTH: 125
TYPE: PRT
ORGANISM: Homo sapiens
US-10-040-862-2292
```

```
Query Match
Best Local Similarity 52.1% Score 61; DB 9; Length 125;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 4 GHGFGNGVPGTGPSCG 20
Db 62 GSGFGSGSGSGSGSGSG 78
```

```
RESULT 10
US-09-796-692-1364
Sequence 1364, Application US/09796692
```

```
Publication No. US20020199362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077,001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,939
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1364
LENGTH: 148
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(148)
OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-1364

Query Match          52.1%; Score 61; DB 9; Length 148;
Best Local Similarity 64.7%; Pred. No. 1.8;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 GHGFGNGVGTGTPGSG 20
DB 99 GSGFGSGSGSGSGSG 115

RESULT 11
US-10-040-862-1364
Sequence 1364, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
```

```
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/223,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1364
LENGTH: 148
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(148)
OTHER INFORMATION: Xaa = Any amino acid
US-10-040-862-1364

Query Match          52.1%; Score 61; DB 9; Length 148;
Best Local Similarity 64.7%; Pred. No. 1.8;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 GHGFGNGVGTGTPGSG 20
DB 99 GSGFGSGSGSGSGSG 115

RESULT 12
US-09-881-752A-118
Sequence 118, Application US/09881752A
Patent No. US20020115078A1
GENERAL INFORMATION:
APPLICANT: Kleantous, Harold
APPLICANT: Al-Garawi, Amal
APPLICANT: Miller, Charles
APPLICANT: Tomb, Jean-Francois
APPLICANT: Oomen, Raymond P.
TITLE OF INVENTION: Identification of Polynucleotides
TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides
FILE REFERENCE: 06132/041002
CURRENT APPLICATION NUMBER: US/09/881,752A
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 08/833,457
PRIOR FILING DATE: 1997-04-01
NUMBER OF SEQ ID NOS: 370
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 118
LENGTH: 154
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-881-752A-118

Query Match          52.1%; Score 61; DB 10; Length 154;
Best Local Similarity 55.0%; Pred. No. 1.9;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

OY 1 NGVHGFGNGVGPCTGPGSG 20
Db 60 GSGFGSGSGSGSGSG 79

RESULT 13

US-09-796-692-2357
Sequence 2357, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND TREATMENT OF
HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2357
LENGTH: 156
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(156)
OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-2357

Query Match 52.1%; Score 61; DB 9; Length 156;
Best Local Similarity 64.7%; Pred. No. 1.9;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 GHGFGNGVGPCTGPGSG 20
Db 93 GSGFGSGSGSGSGSG 109

RESULT 14

US-10-040-862-2357
Sequence 2357, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc

APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and
TITLE OF INVENTION: Hematological Malignancies
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2357
LENGTH: 156
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(156)
OTHER INFORMATION: Xaa = Any amino acid
US-10-040-862-2357

Query Match 52.1%; Score 61; DB 9; Length 156;
Best Local Similarity 64.7%; Pred. No. 1.9;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 GHGFGNGVGPCTGPGSG 20
Db 93 GSGFGSGSGSGSGSG 109

RESULT 15

US-09-925-300-1362
Sequence 1362, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1362
LENGTH: 162
TYPE: PRT

Wed, May 7 14:15:52 2003

us-09-926-234-1.rapb

Page 7

! ORGANISM: Homo sapiens
US-09-925-300-1362

US-09-925-300-1362

Query Match	Score 61;	DB 10;	Length 162;
Post local similarity	64.74	pred No 3.	

Best Local Similarity 64.7%; Pred. No. 2;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Matches	11	Conservative	2	Mismatches	4	Indels	0	Gaps	0
---------	----	--------------	---	------------	---	--------	---	------	---

QY	4	GHFGNGVGPCTGPGSG	20
Dh	99	GSFGSGSGSGSGSGSG	115

99 GSGFGSGSGSGSG 115

Search completed: May 6, 2003, 14:50:54
Job time : 18 secs

Job time : 18 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: May 6, 2003, 14:48:17 ; Search time 15 Seconds
(without alignments)

39.231 Million cell updates/sec

Title: US-09-926-234-1
Sequence: 1 NGVGHSEFNVGPTGPSSG 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum March 0%
Maximum March 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backlist.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	58.1	1169	4	US-08-806-029-33
2	61	52.1	158	1	US-07-906-871-10
3	61	52.1	158	1	US-07-906-871-12
4	60	51.3	1127	4	US-09-150-4608-11
5	59	50.4	142	1	US-07-609-716-100
6	59	50.4	142	4	US-08-475-411A-100
7	59	50.4	142	4	US-08-478-029A-100
8	59	50.4	2018	4	US-09-444-791A-80
9	59	50.4	2100	1	US-08-477-509A-80
10	59	50.4	2100	3	US-09-482-085B-40
11	59	50.4	2107	2	US-08-175-155-45
12	59	50.4	2107	2	US-08-175-155-45
13	58	49.6	93	2	US-08-903-623-43
14	58	49.6	152	1	US-07-906-871-14
15	58	49.6	546	2	US-09-067-351-1
16	58	49.6	546	2	US-09-360-490-1
17	57	48.7	23	1	US-07-609-716-27
18	57	48.7	23	1	US-08-175-155-25
19	57	48.7	23	1	US-08-477-509B-60
20	57	48.7	23	2	US-08-707-237A-31
21	57	48.7	23	3	US-08-482-085B-60
22	57	48.7	23	4	US-08-475-411A-27
23	57	48.7	23	4	US-08-478-029A-27
24	57	48.7	23	4	US-09-444-791A-60
25	57	48.7	23	1	US-08-175-155-8
26	57	48.7	53	1	US-08-477-509B-42
27	57	48.7	53	2	US-09-707-237A-14

28	57	48.7	53	3	US-08-482-085B-42	Sequence 42, App1
29	57	48.7	53	4	US-09-444-791A-42	Sequence 42, App1
30	57	48.7	54	2	US-08-707-237A-56	Sequence 56, App1
31	57	48.7	54	4	US-09-444-791A-116	Sequence 116, App1
32	57	48.7	59	1	US-07-609-716-8	Sequence 8, App1
33	57	48.7	59	1	US-08-175-155-2	Sequence 2, App1
34	57	48.7	59	1	US-08-477-509B-22	Sequence 22, App1
35	57	48.7	59	2	US-08-707-237A-7	Sequence 7, App1
36	57	48.7	59	3	US-08-482-085B-22	Sequence 22, App1
37	57	48.7	59	4	US-08-475-411A-8	Sequence 8, App1
38	57	48.7	59	4	US-08-478-029A-8	Sequence 8, App1
39	57	48.7	59	4	US-09-444-791A-22	Sequence 22, App1
40	57	48.7	63	1	US-07-609-716-98	Sequence 98, App1
41	57	48.7	63	4	US-08-475-411A-98	Sequence 98, App1
42	57	48.7	63	4	US-08-478-029A-98	Sequence 98, App1
43	57	48.7	64	1	US-08-397-633A-55	Sequence 55, App1
44	57	48.7	64	1	US-08-397-633A-56	Sequence 56, App1
45	57	48.7	64	2	US-08-435-641-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-08-806-029-33
Sequence 33, Application US/0806029

Patent No. 6380154
GENERAL INFORMATION:

APPLICANT: Cappello, Joseph
APPLICANT: Stedronsky, Erwin R.
TITLE OF INVENTION: Synthetic Proteins for in vivo Drug
TITLE OF INVENTION: Delivery and Tissue Augmentation
NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:
ADDRESS: Flehr, Hobach, Teet, Albritton & Heathert
STREET: Four Embarradero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1 30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/806,029
FILING DATE: 24-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,237
FILING DATE: 11-MAR-1994

ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOC# NUMBER: A-59847-2/PFT/MTX

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:

LENGTH: 1169 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: protein
US-08-806-029-33

Query Match 58.1% Score 68; DB 4; Length 1169;
Best Local Similarity 73.7%; Pred. No. 0.87;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

2 GVGHSEFNVGPTGPSSG 20

Db 991 GVGFGVGFGAGAGG 1009

RESULT 2

US-07-906-871-10
Sequence 10, Application US/07906871
Patent No. 5340739
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SERGLYCIN AND US8
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W., Suite 300
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,871
FILING DATE: 19920103
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/816,289
FILING DATE: 03 JAN 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/635,544
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/03051
FILING DATE: 13-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/224,035
FILING DATE: 13-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627,2830004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)833-7533
TELEFAX: (202)833-8716
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-906-871-10

Query Match 52.1%; Score 61; DB 1; Length 158;
Best Local Similarity 64.7%; Pred. No. 0.89;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 GHFGNGVGFPTGPGSG 20
DB 95 GSGFGSGSGSGSGSG 111

RESULT 3

US-07-906-871-12
Sequence 12, Application US/07906871
Patent No. 5340739
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.

APPLICANT: Avraham, Shalom
TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SERGLYCIN AND
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W., Suite 300
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,871
FILING DATE: 19920103
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/816,289
FILING DATE: 03 JAN 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/635,544
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/03051
FILING DATE: 13-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/224,035
FILING DATE: 13-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627,2830004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)833-7533
TELEFAX: (202)833-8716
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-906-871-12

Query Match 52.1%; Score 61; DB 1; Length 158;
Best Local Similarity 64.7%; Pred. No. 0.89;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 GHFGNGVGFPTGPGSG 20
DB 95 GSGFGSGSGSGSGSG 111

RESULT 4

US-09-150-460B-11
Sequence 11, Application US/09150460B
Patent No. 6190882
GENERAL INFORMATION:
APPLICANT: Lee, Cheng-Chi
APPLICANT: Albrecht, Ure
APPLICANT: Etchele, Gregor
APPLICANT: Sun, Zhong Sheng
TITLE OF INVENTION: Mammalian Circadian Rhythm-like Gene
FILE REFERENCE: D6039
CURRENT APPLICATION NUMBER: US/09/150,460B
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/054,256
PRIOR FILING DATE: 1997-09-09
NUMBER OF SEQ ID NOS: 21

```

? COUNTRY: US
? ZIP: 94111
? COMPUTER READABLE FORM:
? MEDIUM TYPE: floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/478,029A
? FILING DATE: 07-JUN-1995
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/609,716
? FILING DATE: 06-NOV-1990
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/269,429
? FILING DATE: 09-NOV-1988
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/114,618
? FILING DATE: 29-OCT-1987
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 06/927,258
? FILING DATE: 04-NOV-1986
? ATTORNEY/AGENT INFORMATION:
? NAME: Treacartin, Richard F.
? REGISTRATION NUMBER: 31,801
? REFERENCE/DOCKET NUMBER: A-55196-8/9FT/MTK
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-398-3249
? TELEFAX: 415-781-1989
? INFORMATION FOR SEQ ID NO: 100:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 142 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? US-08-478-029A-100

Query Match 50.4%; Score 59; DB 4; Length 142;
Best Local Similarity 55.0%; Pred. No. 1.4;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 NGVHGFGNGVGPGTGPSSG 20
Db 52 BGAGAGSGAGAGSGAGAGSG 71

RESULT 8
US-09-444-791A-80
? Sequence 80, Application US/09444791A
? Patent No. 6355776
? GENERAL INFORMATION:
? APPLICANT: Ferrari, Franco A.
? APPLICANT: Cappelletto, Joseph
? APPLICANT: Chambers, James
? APPLICANT: Causey, Stuart
? APPLICANT: Collock, Thomas J.
? APPLICANT: Crissman, John W.
? TITLE OF INVENTION: No. 6355776e1 Peptides Comprising Repetitive
? UNITS OF AMINO ACIDS AND DNA SEQUENCES ENCODING THE SAME
? NUMBER OF SEQUENCES: 117
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
? STREET: Four Embarcadero Center, Suite 3400
? CITY: San Francisco
? STATE: California
? COUNTRY: US
? ZIP: 94111
? COMPUTER READABLE FORM:
? MEDIUM TYPE: floppy disk
? COMPUTER: IBM PC compatible

```

```

? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/444,791A
? FILING DATE: 22-NOV-6355776-1999
? CLASSIFICATION: <Unknown>
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/482,085
? FILING DATE: 07-JUN-1995
? APPLICATION NUMBER: US 08/175,155
? FILING DATE: 29-DEC-1993
? APPLICATION NUMBER: US 08/053,049
? FILING DATE: 22-APR-1993
? APPLICATION NUMBER: US 07/114,618
? FILING DATE: 29-OCT-1987
? APPLICATION NUMBER: US 06/927,258
? FILING DATE: 04-NOV-1986
? ATTORNEY/AGENT INFORMATION:
? NAME: Treacartin, Richard F.
? REGISTRATION NUMBER: 31,801
? REFERENCE/DOCKET NUMBER: A-55196-11/PFT/BTC
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-398-3249
? TELEFAX: 415-781-1989
? INFORMATION FOR SEQ ID NO: 80:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2018 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? SEQUENCE DESCRIPTION: SEQ ID NO: 80:
US-09-444-791A-80

Query Match 50.4%; Score 59; DB 4; Length 2018;
Best Local Similarity 57.9%; Pred. No. 18;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 GVGHGFGNGVGPGTGPSSG 20
Db 1982 GVAAGYGAGAGSGAGAGSG 2000

RESULT 9
US-08-477-509B-80
? Sequence 80, Application US/08477509B
? Patent No. 5770697
? GENERAL INFORMATION:
? APPLICANT: Ferrari, Franco A.
? APPLICANT: Cappelletto, Joseph
? APPLICANT: Crissman, John W.
? APPLICANT: Dorman, Mary A.
? TITLE OF INVENTION: No. 5770697e1 Peptides Comprising Repetitive
? UNITS OF AMINO ACIDS AND DNA SEQUENCES ENCODING THE S
? NUMBER OF SEQUENCES: 112
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
? STREET: Four Embarcadero Center, Suite 3400
? CITY: San Francisco
? STATE: California
? COUNTRY: US
? ZIP: 94111
? COMPUTER READABLE FORM:
? MEDIUM TYPE: floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/477,509B
? FILING DATE: 07-JUN-1995
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/175,155

```

FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,045
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-7/PFT/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 2100 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-477-509B-80

Query Match 50.4%; Score 59; DB 1; Length 2100;
Best Local Similarity 57.9%; Pred. No. 19;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 GVHGFGNGVPGTGPSSG 20
DB 2064 GVAAGYGAGSGAGSG 2082

RESULT 10
US-08-482-085B-80
Sequence 80, Application US/08482085B
Patent No. 6018030
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Richardson, Charles
APPLICANT: Chambers, James
APPLICANT: Causey, Stuart
APPLICANT: Pollock, Thomas J.
APPLICANT: Cappello, Joseph
APPLICANT: Crissman, John W.
TITLE OF INVENTION: No. 6018030el Peptides Comprising Repetitive
TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
NUMBER OF SEQUENCES: 112
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,085B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/053,045
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,155
FILING DATE: 29-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-6/PFT/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 2100 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-482-085B-80

Query Match 50.4%; Score 59; DB 3; Length 2100;
Best Local Similarity 57.9%; Pred. No. 19;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 GVHGFGNGVPGTGPSSG 20
DB 2064 GVAAGYGAGSGAGSG 2082

RESULT 11
US-08-175-155-45
Sequence 45, Application US/08175155
Patent No. 5641648
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Cappello, Joseph
APPLICANT: Crissman, John W.
APPLICANT: Dorman, Mary A.
TITLE OF INVENTION: Methods for Preparing Synthetic
TITLE OF INVENTION: Repetitive DNA
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,155
FILING DATE: 29-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I.
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: A-55186-5/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 2107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-175-155-45

US-08-707-237A-51

US-07-906-871-14

US-07-906-871-14

Patent No. 5340739
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC
TRANSCRIPTIONAL REGULATORY ELEMENTS OF SERGLYCIN AND USES
THEREOF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W., Suite 300
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM: **Y**
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1 0, Version #1 25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,871
FILING DATE: 19920103
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/816,289
FILING DATE: 03 JAN 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/635,544
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/03051
FILING DATE: 13-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/224,035
FILING DATE: 13-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627,2830004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)833-7533
TELEFAX: (202)833-8716
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-906-871-14

Query Match 49.6%; Score 58; DB 1; Length 152;
Best Local Similarity 55.0%; Pred. NO. 2;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 NGVGHFGNGVGPGTGPGSG 20
DB 90 SSGSGSGSGSGSGSGSGSG 109

RESULT 15
US-09-067-351-1
Sequence 1, Application US/09067351
Patent No. 5994081
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: HUMAN KERATINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/067,351
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0511 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 546 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PANCYT02
CLONE: 1467090
US-09-067-351-1

Query Match 49.6%; Score 58; DB 2; Length 546;
Best Local Similarity 57.9%; Pred. NO. 6.7;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 GVGHFGNGVGPGTGPGSG 20
DB 92 GGGYGFGGAGSGGFGGCG 110

Search completed: May 6, 2003, 14:50:31
Job time : 16 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on:

May 6, 2003, 14:44:02 ; Search time 35 seconds

(without alignments)

76.143 Million cell updates/sec

Title: US-09-926-234-1

Perfect score: 117

Sequence: 1 NGVGHFGNGVGPGRGPGSG 20

Scoring table:

BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched:

908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:

908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /SID2/gcgdata/geneeq/geneeq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneeq/geneeq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneeq/geneeq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneeq/geneeq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneeq/geneeq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneeq/geneeq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneeq/geneeq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneeq/geneeq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneeq/geneeq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneeq/geneeq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneeq/geneeq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneeq/geneeq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneeq/geneeq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneeq/geneeq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneeq/geneeq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneeq/geneeq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneeq/geneeq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneeq/geneeq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneeq/geneeq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneeq/geneeq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneeq/geneeq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneeq/geneeq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneeq/geneeq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	100.0	20	21	AAB26878
2	80	68.4	270	22	ABB59223
3	71	60.7	1260	22	ABB59060
4	69	59.0	206	21	AAAG25348
5	69	59.0	221	21	AAAG26579
6	69	59.0	239	21	AAAG26578
7	69	59.0	296	21	AAAG26577
8	69	59.0	316	21	AAAG4983
9	69	59.0	384	21	AAAG37993
10	69	59.0	402	21	AAAG43303

11	69	59.0	416	21	AAAG37892	Arabidopsis thaliana
12	69	59.0	427	21	AAAG37891	Arabidopsis thaliana
13	69	59.0	434	21	AAAG4402	Arabidopsis thaliana
14	69	59.0	444	21	AAAG4401	Arabidopsis thaliana
15	68	58.1	20	21	AAB26879	Glycine rich peptide
16	68	58.1	135	21	AAAG25286	Arabidopsis thaliana
17	68	58.1	148	21	AAAG25285	Arabidopsis thaliana
18	64	54.7	372	22	ABB69337	Arabidopsis thaliana
19	64	54.7	372	22	ABB69337	Arabidopsis thaliana
20	63	53.8	40	23	AAO18249	Arabidopsis thaliana
21	63	53.8	84	23	AAO18248	Arabidopsis thaliana
22	63	53.8	104	22	AAU56019	Arabidopsis thaliana
23	63	53.8	104	22	AAU56019	Arabidopsis thaliana
24	63	53.8	687	22	AAO18251	Arabidopsis thaliana
25	63	53.8	2703	22	ABB63299	Arabidopsis thaliana
26	62	53.0	60	22	AAU64705	Arabidopsis thaliana
27	62	53.0	219	23	AAU50040	Arabidopsis thaliana
28	62	53.0	264	23	AAU50040	Arabidopsis thaliana
29	62	53.0	285	22	ABB65204	Arabidopsis thaliana
30	62	53.0	294	22	ABB65204	Arabidopsis thaliana
31	62	53.0	319	19	AAW72059	Arabidopsis thaliana
32	62	53.0	340	19	AAW72149	Arabidopsis thaliana
33	62	53.0	340	19	AAW72149	Arabidopsis thaliana
34	61	52.1	59	22	AAW80990	Arabidopsis thaliana
35	61	52.1	59	22	AAW80990	Arabidopsis thaliana
36	61	52.1	88	21	AAAG1556	Arabidopsis thaliana
37	61	52.1	105	21	AAAG1556	Arabidopsis thaliana
38	61	52.1	105	21	AAAG1556	Arabidopsis thaliana
39	61	52.1	107	21	AAAG1556	Arabidopsis thaliana
40	61	52.1	117	22	AAU44209	Arabidopsis thaliana
41	61	52.1	125	22	AAU44209	Arabidopsis thaliana
42	61	52.1	125	22	AAU44209	Arabidopsis thaliana
43	61	52.1	125	22	AAU44209	Arabidopsis thaliana
44	61	52.1	134	21	AAAG20185	Arabidopsis thaliana
45	61	52.1	134	21	AAAG20185	Arabidopsis thaliana
			148	22	AAW81000	Human haematologic

ALIGNMENTS

RESULT 1
ID AAB26878 standard; peptide; 20 AA.
AC AAB26878;
DT 01-FEB-2001 (first entry)
DE Glycine rich peptide SEQ ID 1 for treating multiple sclerosis.
KW Immunosuppressant; neuroprotective; multiple sclerosis; treatment.
OS Synthetic.
XX WO200058354-A1.
XX PN 05-OCT-2000.
XX PD 29-MAR-1999; 99WO-EP02268.
XX PF 29-MAR-1999; 99WO-EP02268.
XX PR 29-MAR-1999; 99WO-EP02268.
XX PA (TECN-) TECHNICAL SCPA.
XX PI Marino M, Ippolito A, Fassina G;
XX WPI; 2000-628344/60.
XX DR New glycine-rich peptides, useful for treatment of multiple sclerosis
XX PT by induction of energy in autoreactive T cells
XX PS Claim 1; Page 13; 23pp, English.

CC Peptides ABB6878-B26881 represent four glycine rich peptide which may
 CC be N-acetylated and/or C-extended and contain amino acids with L or D
 CC configuration. Included in the invention is a pharmaceutical composition
 CC containing at least one of the peptides and an inert ingredient. The
 CC peptides have immunosuppressant and neuroprotective activity and induce
 CC energy of autoreactive T lymphocytes without inducing an autoreactive
 CC response. The peptides are used to treat multiple sclerosis.

XX Sequence 20 AA;

Query Match

Best Local Similarity 100.0%; Score 117; DB 21; Length 20;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVGPGTGPSSG 20

Db 1 NGVGHGFGNGVGPGTGPSSG 20

RESULT 2

ABBS9223

ID ABB59223 standard; Protein; 270 AA.

XX ABB59223;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 4461

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN W0200171042-A2

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI: 2001-656860/75

DR N-PSDB; ABL03163.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 4461; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 270 AA;

Query Match

Best Local Similarity 68.4%; Score 80; DB 22; Length 270;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVGPGTGPSSG 20

Db 210 NGVGHGFGNGVGPGTGPSSG 229

RESULT 3

ABBS9060

ID ABB59060 standard; Protein; 1260 AA.

XX ABB59060;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 3972.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN W0200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PP 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI: 2001-656860/75

DR N-PSDB; ABL03163.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 3972; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1260 AA;

Query Match

Best Local Similarity 60.7%; Score 71; DB 22; Length 1260;

Matches 14; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 NGVGHGFGNGVGPGTGPSSG 20

Db 1019 NGVGHGFGNGVGPGTGPSSG 1038

RESULT 4

AAG25348

ID AAG25348 standard; Protein; 206 AA.

XX AAG25348;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29375.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
PD
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-011825.
PR 05-MAR-1999; 99US-0121180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 26-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142970.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149476.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153768.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155658.
PR 28-SEP-1999; 99US-0157158.
PR 29-SEP-1999; 99US-0158196.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157751.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match# 59.0%; Score 69; DB 21; Length 206;
Best Local Similarity 64.7%; Pred. No. 0.64;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 GHGFGNGVPGCTGPGSG 20
Db 93 GSGFGGIGGSGGPGSG 109

RESULT 5
AAG26579
ID AAG26579 standard; Protein: 221 AA

XX AAG26579;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO. 31087

XX Protein identification, signal transduction pathway, metabolic pathway,
XX Hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX

OS Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0129645.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135253.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137232.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139464.

PR 18-JUN-1999; 99US-0139465.

PR 18-JUN-1999; 99US-0139466.

PR 18-JUN-1999; 99US-0139467.

PR 18-JUN-1999; 99US-0139468.

PR 18-JUN-1999; 99US-0139469.

PR 18-JUN-1999; 99US-0139470.

PR 18-JUN-1999; 99US-0139471.

PR 18-JUN-1999; 99US-0139472.

PR 18-JUN-1999; 99US-0139473.

PR 18-JUN-1999; 99US-0139474.

PR 18-JUN-1999; 99US-0139475.

PR 18-JUN-1999; 99US-0139476.

PR 18-JUN-1999; 99US-0139477.

PR 18-JUN-1999; 99US-0139478.

PR 18-JUN-1999; 99US-0139479.

PR 18-JUN-1999; 99US-0139480.

PR 18-JUN-1999; 99US-0139481.

PR 18-JUN-1999; 99US-0139482.

PR 18-JUN-1999; 99US-0139483.

PR 18-JUN-1999; 99US-0139484.

PR 18-JUN-1999; 99US-0139485.

PR 18-JUN-1999; 99US-0139486.

PR 18-JUN-1999; 99US-0139487.

PR 18-JUN-1999; 99US-0139488.

PR 18-JUN-1999; 99US-0139489.

PR 18-JUN-1999; 99US-0139490.

PR 18-JUN-1999; 99US-0139491.

PR 18-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139493.

PR 18-JUN-1999; 99US-0139494.

PR 18-JUN-1999; 99US-0139495.

PR 18-JUN-1999; 99US-0139496.

PR 18-JUN-1999; 99US-0139497.

PR 18-JUN-1999; 99US-0139498.

PR 18-JUN-1999; 99US-0139499.

PR 18-JUN-1999; 99US-0139500.

PR 18-JUN-1999; 99US-0139501.

PR 18-JUN-1999; 99US-0139502.

PR 18-JUN-1999; 99US-0139503.

PR 18-JUN-1999; 99US-0139504.

PR 18-JUN-1999; 99US-0139505.

PR 18-JUN-1999; 99US-0139506.

PR 18-JUN-1999; 99US-0139507.

PR 18-JUN-1999; 99US-0139508.

PR 18-JUN-1999; 99US-0139509.

PR 18-JUN-1999; 99US-0139510.

PR 18-JUN-1999; 99US-0139511.

PR 18-JUN-1999; 99US-0139512.

PR 18-JUN-1999; 99US-0139513.

PR 18-JUN-1999; 99US-0139514.

PR 18-JUN-1999; 99US-0139515.

PR 18-JUN-1999; 99US-0139516.

PR 18-JUN-1999; 99US-0139517.

PR 18-JUN-1999; 99US-0139518.

PR 18-JUN-1999; 99US-0139519.

PR 18-JUN-1999; 99US-0139520.

PR 18-JUN-1999; 99US-0139521.

PR 18-JUN-1999; 99US-0139522.

PR 18-JUN-1999; 99US-0139523.

PR 18-JUN-1999; 99US-0139524.

PR 18-JUN-1999; 99US-0139525.

PR 18-JUN-1999; 99US-0139526.

PR 18-JUN-1999; 99US-0139527.

PR 18-JUN-1999; 99US-0139528.

PR 18-JUN-1999; 99US-0139529.

PR 18-JUN-1999; 99US-0139530.

PR 18-JUN-1999; 99US-0139531.

PR 18-JUN-1999; 99US-0139532.

PR 18-JUN-1999; 99US-0139533.

PR 18-JUN-1999; 99US-0139534.

PR 18-JUN-1999; 99US-0139535.

PR 18-JUN-1999; 99US-0139536.

PR 18-JUN-1999; 99US-0139537.

PR 18-JUN-1999; 99US-0139538.

PR 18-JUN-1999; 99US-0139539.

PR 18-JUN-1999; 99US-0139540.

PR 18-JUN-1999; 99US-0139541.

PR 18-JUN-1999; 99US-0139542.

PR 18-JUN-1999; 99US-0139543.

PR 18-JUN-1999; 99US-0139544.

PR 18-JUN-1999; 99US-0139545.

PR 18-JUN-1999; 99US-0139546.

PR 18-JUN-1999; 99US-0139547.

PR 18-JUN-1999; 99US-0139548.

PR 18-JUN-1999; 99US-0139549.

PR 18-JUN-1999; 99US-0139550.

PR 18-JUN-1999; 99US-0139551.

PR 18-JUN-1999; 99US-0139552.

PR 18-JUN-1999; 99US-0139553.

PR 18-JUN-1999; 99US-0139554.

PR 18-JUN-1999; 99US-0139555.

PR 18-JUN-1999; 99US-0139556.

PR 18-JUN-1999; 99US-0139557.

PR 18-JUN-1999; 99US-0139558.

PR 18-JUN-1999; 99US-0139559.

PR 18-JUN-1999; 99US-0139560.

PR 18-JUN-1999; 99US-0139561.

PR 18-JUN-1999; 99US-0139562.

PR 18-JUN-1999; 99US-0139563.

PR 18-JUN-1999; 99US-0139564.

PR 18-JUN-1999; 99US-0139565.

PR 18-JUN-1999; 99US-0139566.

PR 18-JUN-1999; 99US-0139567.

PR 18-JUN-1999; 99US-0139568.

PR 18-JUN-1999; 99US-0139569.

PR 18-JUN-1999; 99US-0139570.

PR 18-JUN-1999; 99US-0139571.

PR 18-JUN-1999; 99US-0139572.

PR 18-JUN-1999; 99US-0139573.

PR 18-JUN-1999; 99US-0139574.

PR 18-JUN-1999; 99US-0139575.

PR 18-JUN-1999; 99US-0139576.

PR 18-JUN-1999; 99US-0139577.

PR 18-JUN-1999; 99US-0139578.

PR 18-JUN-1999; 99US-0139579.

PR 18-JUN-1999; 99US-0139580.

PR 18-JUN-1999; 99US-0139581.

PR 18-JUN-1999; 99US-0139582.

PR 18-JUN-1999; 99US-0139583.

PR 18-JUN-1999; 99US-0139584.

PR 18-JUN-1999; 99US-0139585.

PR 18-JUN-1999; 99US-0139586.

PR 18-JUN-1999; 99US-0139587.

PR 18-JUN-1999; 99US-0139588.

PR 18-JUN-1999; 99US-0139589.

PR 18-JUN-1999; 99US-0139590.

PR 18-JUN-1999; 99US-0139591.

PR 18-JUN-1999; 99US-0139592.

PR 18-JUN-1999; 99US-0139593.

PR 18-JUN-1999; 99US-0139594.

PR 18-JUN-1999; 99US-0139595.

PR 18-JUN-1999; 99US-0139596.

PR 18-JUN-1999; 99US-0139597.

PR 18-JUN-1999; 99US-0139598.

PR 18-JUN-1999; 99US-0139599

```

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143622.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145226.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147018.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

```

```

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158212.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162182.

```

Query Match 59.0%; Score 69; DB 21; Length 221;
 Best Local Similarity 64.7%; Pred. No. 0.68; 4; Indels 0; Gaps 0;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

CY 4 GGGGNGVPGTGGG 20
 DB 173 GGGGNGVPGTGGG 189

```

RESULT 6
AACG26578
ID AACG26578 standard; Protein; 239 AA.
XX
AC AACG26578;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 31086.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.

```

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130710.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134238.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134994.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147208.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158366.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159738.
PR 18-OCT-1999; 99US-0159884.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.0%; Score 69; DB 21; Length 239;
Best Local Similarity 64.7%; Pred. No. 0.73;

Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 GHGFGNGVPGTGTGSG 20
| | | | | | | | | |
Dh 191 GCGFGGIGGSGGPGEG 207

RESULT 7

AAG26577 standard, Protein, 296 AA.

XX AAG26577;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 31085.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.

OS EP1033405-A2.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301419

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132683.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134378.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132864.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0135141.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141844.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143544.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148568.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158265.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.0%; Score 69; DB 21; Length 384;
Best Local Similarity 64.7%; Pred. No. 1.1;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 GHGFGNGVGPCTGPSSG 20
DB 271 GSGFGEGTGGSGPDEG 287

RESULT 10
ID AAG44303 standard; Protein; 402 AA.
XX AAG44303;
AC AAG44303;
XX 18-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55479.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128214.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135114.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137328.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138410.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	17-JUN-1999;	99US-0139453.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141281.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142399.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144088.
PR	16-JUL-1999;	99US-0144088.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145089.
PR	23-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147392.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147703.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147933.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149215.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	22-SEP	

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 59.0%; Score 69; DB 21; Length 402;
Best Local Similarity 64.7%; Pred. No. 1.2;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 CHGFGNGVGTGTGPGSG 20
DB 289 GSGFGGTCGSGGPGGCG 705

RESULT 11
ID AAG37892 standard; Protein; 416 AA.
XX
AC AAG37892;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46667.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence
XX
XX Arabidopsis thaliana.
XX OS
XX PN EPI03405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-012180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 24-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132488.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140364.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142970.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145961.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147320.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149726.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.0%; Score 69; DB 21; Length 416;
Best Local Similarity 64.7%; Pred. No. 1.2;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 4 GGGFGNGVGPOTGPGSG 20
Db 303 GSGFGGIGGSGGPGG 319
RESULT 12
AAG37891
ID AAG37891 standard; Protein; 427 AA.
XX
AC AAG37891;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4666.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydropathicity assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126745.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134378.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.

```
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139465.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.0%; Score 69; DB 21; Length 427;
Best Local Similarity 64.7%; Pred No. 1.2;
Matches 11; Conservative 2; Mismatches 0; Gaps 0;
QY 4 GHFGNGVPGTGGSG 20
|||||:|:|:|
```

DB 314 GSGFEGIGGSGGEG 330

RESULT 13

AAG44302

ID AAG44302 standard: Protein: 434 AA.

XX AAG44302;

AC AAG44302;

XX 18-OCT-2000 (first entry)

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55478.

XX Protein identification: signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

OS Arabidopsis thaliana.

PN EPI03405-A2.

XX 06-SEP-2000.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0101439

PF 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 14-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139464.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147312.

PR 05-AUG-1999; 99US-0147360.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0151963.
PR 10-SEP-1999; 99US-0152070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158237.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.0%; Score 69; DB 21; Length 434;
Best Local Similarity 64.7%; Pred. No. 1.3;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 CHGFGNAGVPTGTPGSG 20
DB 321 GSGFGEGIGGSGSPGEG 337

RESULT 14
AAC44301
ID AAC44301 standard; Protein; 444 AA.
XX

AC AAC44301;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55477.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136032.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

```

PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139761.
PR 21-JUN-1999; 99US-0139813.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140351.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140821.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142155.
PR 06-JUL-1999; 99US-0142190.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144633.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145127.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

```

```

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157855.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 59.04; Score 69; DB 21; Length 444;
 Best Local Similarity 64.74; Pred. No. 13;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 GHFGNGVCPGTGPGSG 20
 DB 331 GSGFGGFGTGGSGGPGG 347

RESULT 15
 AAB26879 standard; peptide: 20 AA.

XX AAB26879;

DT 01-FEB-2001 (first entry)

DE Glycine rich peptide SEQ ID 2 for treating multiple sclerosis.

KW Immunosuppressant; neuroprotective; multiple sclerosis; treatment.

```

XX OS Synthetic.
XX PN WO20058354-A1.
XX PD 05-OCT-2000.
XX PF 29-MAR-1999; 99WO-EP02268.
XX PR 29-MAR-1999; 99WO-FP02268.
XX PA (TECN.) TECHNOGEN SCPA.
XX PI Marino M, Ippolito A, Fassina G.
XX DR WPI; 2000-628344/60.
XX PT New glycine-rich peptides, useful for treatment of multiple sclerosis
XX PT by induction of anergy in autoreactive T cells -
XX PS Claim 1; Page 13; 23pp; English.
XX CC Peptides AAB26878-B26881 represent four glycine rich peptide which may
XX CC be N-acetylated and/or C-amidated and contain amino acids with L or D
XX CC configuration. Included in the invention is a pharmaceutical composition
XX CC containing at least one of the peptides and an inert ingredient. The
XX CC peptides have immunosuppressant and neuroprotective activity and induce
XX CC anergy of autoreactive T lymphocytes without inducing an autoreactive
XX CC response. The peptides are used to treat multiple sclerosis.
SQ Sequence 20 AA:

Query Match 58.1%; Score 68; DB 21; Length 20;
Best Local Similarity 63.2%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVGHGFGNGVGPGTGPQSG 20
   | | | | | | | | | |
Db 1 GPGVGFGKGIOTGRGPG 19

```

Search completed: May 6, 2003, 14:48:55
 Job time : 36 secs